General Internal Medicine
Boston University School of Medicine
2012 Publications A-K

1. Ageno W, Gallus AS, Wittkowsky A, Crowther M, **Hylek EM**, Palareti G. Oral anticoagulant therapy: Antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. 2012 Feb; 141(Suppl 2):e44S-88S. PMID: 22315269.

2. Allensworth-Davies D, **Cheng DM**, **Smith PC**, **Samet JH**, **Saitz R**. The Short Inventory of Problems-modified for Drug Use (SIP-DU): Validity in a primary care sample. Am J Addict. 2012 May-Jun;21(3):257-62. PMID: 22494228.

3. Ayotte BJ, Hausmann LR, Whittle J, **Kressin NR**. The relationship between perceived discrimination and coronary artery obstruction. Am Heart J. 2012 Apr; 163(4):677-83. PMID: 22520534.

4. Bailey SC, McCormack LA, Rush SR, **Paasche-Orlow** MK Advancing health literacy research. J Health Commun. 2012; 17(Suppl 3):2-6. PMID: 23030556.

5. **Battaglia TA**, Howard MB, Kavanah M, Prout MN, **Chapman C**, **David MM**, **McKinney R**, **Kronman A**, **Dumont T**, **Freund KM**. The addition of internists to a breast health program. Breast J. 2012 Jan-Feb; 18(1):58-64. doi: 10.1111/j.1524-4741.2011.01179.x. PMID: 22098389.

6. **Battaglia TA**, McCloskey L, **Caron SE**, Murrell SS, Bernstein E, Childs A, Jong H, Walker K, Bernstein J. Feasibility of chronic disease patient navigation in an urban primary care practice. J Ambul Care Manage. 2012 Jan;35(1):38-49. PMID: 22156954.

7. **Battaglia TA**, Murrell SS, Bhosrekar SG, Caron SE, Bowen DJ, Smith E, Kalish R, Rorie JA. Connecting Boston's public housing developments to community health centers: Who's ready for change? Prog Community Health Partnersh. 2012 Fall;6(3):239-48. PMID: 22982838.

8. **Battaglia TA**, **Bak SM**, Heeren T, Chen CA, **Kalish R**, Tringale S, Taylor JO, Lottero B, Egan AP, Thakrar N, **Freund KM**. Boston patient navigation research program: The impact of navigation on time to diagnostic resolution after abnormal cancer screening. Cancer Epidemiol Biomarkers Prev. 2012 Oct;21(10):1645-54. Doi:10.1158/1055-9965.EPI-12-0532. PMID: 23045539.

9. Beletsky L, Rich JD, **Walley AY**. Prevention of fatal opioid overdose. JAMA. 2012 Nov;308(18):1863-64. PMID: 23150005.

10. Berkowitz RE, Schreiber R, **Paasche-Orlow MK**. Team improvement and patient safety conferences: Culture change and slowing the revolving door between skilled nursing facility and the hospital. J Nurs Care Qual. 2012 Jul; 27(3):258-65. PMID: 22361932.

11. Bickmore TW, **Paasche-Orlow** MK. The role of information technology in health literacy research. J Health Commun. 2012; 17(Suppl 3):23-9. PMID: 23030559.

12. Blood EA, **Cheng DM**. Non-linear mixed models in the analysis of mediated longitudinal data with binary outcomes. BMC Med Res Methodol. 2012 Jan; 12(1):5. PMID: 22273051.

13. Bokhour BG, Cohn ES, Cortés DE, Solomon JL, Fix GM, Elwy AR, Mueller N, Katz LA, Haidet P, Green AR, **Borzecki AM**, **Kressin NR**. The role of patients' explanatory models and daily-lived experience in hypertension self-management. J Gen Intern Med. 2012 Dec; 27(12):1626-34. PMID: 22821569.

14. **Borzecki AM,** Cowan AJ, Cevasco M, Shin MH, Shwartz M, Itani K, Rosen AK. Is development of postoperative venous thromboembolism related to thromboprophylaxis use? A case-control study in the Veterans Health Administration. Jt Comm J Qual Patient Saf. 2012 Aug; 38(8):348-58. PMID: 22946252.

15. Bottonari KA, Tripathi SP, Fortney JC, Curran G, Rimland D, Rodriguez-Barradas M, **Gifford AL**, Pyne JM. Correlates of antiretroviral and antidepressant adherence among depressed HIV-infected patients. AIDS Patient Care STDS. 2012 May; 26(5):265-73. PMID: 22536930.

16. Britton KA, Pedley A, Massaro JM, Corsini EM, **Murabito JM**, Hoffmann U, Fox CS Prevalence, distribution, and risk factor correlates of high thoracic periaortic fat in the Framingham Heart Study. Am Heart Assoc. 2012 Dec;1(6):e004200. PMID: 23316328.

17. Brooks DR, Avetisyan R, Jarrett KM, **Hanchate A**, Shapiro GD, Pugh MJ, **Berlowitz DR**, Thurman D, Montouris G, Kazis LE. Validation of self-reported epilepsy for purposes of community surveillance. Epilepsy Behav. 2012 Jan; 23(1):57-63. PMID: 22189155.

18. Cohen MJ, Shaykevich S, Cawthon C, Kripalani S, **Paasche-Orlow MK**, Schnipper JL. Predictors of medication adherence postdischarge: The impact of patient age, insurance status, and prior adherence. J Hosp Med. 2012 Jul-Aug; 7(6):470-5. PMID: 22473754. doi: 10.1002/jhm.1940.

19. **Crosby SS**. A doctor's response to torture. Ann Intern Med. 2012 Mar;156(6):471-2. PMID: 22431681.

20. Daley JI, Stahre MA, Chaloupka FJ, **Naimi TS**. The impact of a 25-cent-per-drink alcohol tax increase. Am J Prev Med. 2012 Apr; 42(4):382-9. PMID: 22424251.

21. **D Amore MM**, **Cheng DM**, Allensworth-Davies D, **Samet JH**, **Saitz R**. Disparities in safe sex counseling & behavior among individuals with substance dependence: A cross-sectional study. Reprod Health. 2012 Dec;9(1):35. PMID: 23276300.

22. Drainoni ML, Litwin AH, Smith BD, Koppelman EA, McKee MD, Christiansen CL, **Gifford AL**, Weinbaum CM, Southern WN. Effectiveness of a risk screener in identifying hepatitis C virus in a primary care setting. Am J Public Health. 2012 Nov; 102(11):e115-21. PMID: 22994166.

23. Driver JA, Beiser A, Au R, **Kreger BE**, Splansky GL, Kurth T, Kiel DP, Lu KP, Seshadri S, Wolf PA. Inverse association between cancer and Alzheimer's disease: Results from the Framingham Heart Study. BMJ. 2012 Mar; 344:e1442. doi: 10.1136/bmj.e1442. PMID: 22411920.

24. Esser MB, Kanny D, Brewer RD, **Naimi TS**. Binge drinking intensity: A comparison of two measures. Am J Prev Med. 2012 Jun; 42(6):625-9. PMID: 22608381.

25. Evans JL, **Tsui JI**, Hahn JA, Davidson PJ, Lum PJ, Page K. Mortality among young injection drug users in San Francisco: A 10-year Follow-up of the UFO Study. Am J Epidemiol. 2012 Feb; 175(4):302-8. PMID: 22227793.

26. **Fleurant M**, Kell R, Jenter C, Volk LA, Zhang F, Bates DW, Simon SR. Factors associated with difficult electronic health record implementation in office practice. J Am Med Inform Assoc. 2012; 19:541-44. doi:10.1136/amiajnl-2011-000689. PMID: 22249965.

27. **Gerber MR**, Fried LE, Pineles SL, Shipherd JC, Bernstein CA. Posttraumatic stress disorder and intimate partner violence in a women's headache center. Women Health. 2012 Jul; 52(5):454-71. PMID: 22747183

28. Gidwani R, Goetz MB, Kominski G, Asch S, Mattocks K, **Samet JH**, Justice A, Gandhi N, Needleman J. A budget impact analysis of rapid human immunodeficiency virus screening in Veterans Administration emergency departments. J Emerg Med. 2012 Jun;42(6):719-26. PMID: 21277144.

29. Gordon AJ, **Alford DP**. Screening, Brief Intervention, and Referral to Treatment (SBIRT) curricular innovations: Addressing a training gap. Subst Abus. 2012 Jul; 33(3):227-30. PMID: 22737999.

30. Groessl EJ, Liu L, Ho SB, Kanwal F, **Gifford AL**, Asch SM. National patterns and predictors of liver biopsy use for management of hepatitis C. J Hepatol. 2012 Aug; 57(2):252-9. PMID: 22521358.

31. Hahn RA, Middleton JC, Elder R, Brewer R, Fielding J, **Naimi TS**, Toomey TL, Chattopadhyay S, Lawrence B, Campbell CA; Community Preventive Services Task Force. Effects of alcohol retail privatization on excessive alcohol consumption and related harms: A community guide systematic review. Am J Prev Med. 2012 Apr; 42(4):418-27. PMID: 22424256.

32. **Hanchate AD**, **Lasser KE**, **Kapoor A**, Rosen J, McCormick D, **D'Amore MM**, **Kressin NR**. Massachusetts reform and disparities in inpatient care utilization. Med Care. 2012 Jul; 50(7):569-77. PMID: 22683590.

33. **Hanchate AD**, Schwamm LH, Huang W, **Hylek EM.** Comparison of ischemic stroke outcomes and patient and hospital characteristics by race/ethnicity and socioeconomic status. Stroke. 2013 Feb; 44(2):469-76. PMID: 23306327.

34. Haring R, Teng Z, Xanthakis V, Coviello A, Sullivan L, Bhasin S, **Murabito JM,** Wallaschofski H, Vasan RS. Association of sex steroids, gonadotrophins, and their trajectories with clinical cardiovascular disease and all-cause mortality in elderly men from the Framingham Heart Study. Clin Endocrinol (Oxf). 2012 Aug ;(78):629-34. PMID: 22901104.

35. Haring R, Xanthakis V, Coviello A, Sullivan L, Bhasin S, Wallaschofski H, **Murabito JM**, Vasan RS. Clinical correlates of sex steroids and gonadotropins in men over the late adulthood: The Framingham Heart Study. Int J Androl. 2012. Dec; 35(6):775-82. PMID: 22640232.

36. He C, Chasman DI, Dreyfus J, Hwang SJ, Ruiter R, Sanna S, Buring JE, Fernández-Rhodes L, Franceschini N, Hankinson SE, Hofman A, Lunetta KL, Palmieri G, Porcu E, Rivadeneira F, Rose LM, Splansky GL, Stolk L, Uitterlinden AG, Chanock SJ, Crisponi L, Demerath EW, **Murabito JM**, Ridker PM, Stricker BH, Hunter DJ. Reproductive aging-associated common genetic variants and the risk of breast cancer. Breast Cancer Res. 2012 Mar 20; 14(2):R54. PMID: 22433456.

37. **Hermos JA**, Young, MM, Fonda J, Gagnon DR, Fiore LD, Lawler EV. Risk of community-acquired pneumonia in veteran patients dispensed proton pump inhibitors. Clinical Infectious Diseases. 2012 Jan;54(1):33-42. PMID: 22100573.

38. Hlatky MA, Douglas PS, Cook NL, Wells B, Benjamin EJ, Dickersin K, Goff DC, Hirsch AT, **Hylek EM**, Peterson ED, Roger VL, Selby JV, Udelson JE, Lauer MS. Future directions for cardiovascular disease comparative effectiveness research: Report of a workshop sponsored by the National Heart, Lung, and Blood Institute. J Am Coll Cardiol. 2012; 2012 Aug 14;60(7):569-80. PMID: 22796257.

39. Howard MB, **Battaglia T**, Prout M, Freund K. The effect of imaging on the clinical management of breast pain. J Gen Intern Med. 2012 Jul; 27(7):817-24 PMID: 22331398.

40. Jasuja GK, Reisman JI, Miller DR, **Berlowitz DR, Hylek EM,** Ash AS, Ozonoff A, Zhao S, **Rose AJ.** Identifying major hemorrhage with automated data: Results of the Veterans Affairs Study to improve anticoagulation (VARIA). Thromb Res. 2013 Jan; 131(1):31-6. PMID: 23158402.

41. Kemppainen J, Bormann JE, Shively M, Kelly A, Becker S, Bone P, Belding W, **Gifford AL**. Living with HIV: Responses to a mantram intervention using the critical incident research method. Altern Complement Med. 2012 Jan; 18(1):76-82PMID: 22268972.

42. **Kim TW, Saitz R, Cheng DM,** Winter MR, Witas J, **Samet JH.** Effect of quality chronic disease management for alcohol and drug dependence on addiction outcomes. J Subst Abuse Treat. 2012 Dec; 43(4):389-96. PMID: 22840687.

43. **Kim TW, Samet JH.** Co-morbidity is the norm, not the exception: Chronic respiratory diseases in chronic drug users. Prim Care Respir J 2012 Dec; 21(4):364-5. PMID: 23138843.

44. **Kronman AC**, **Freund KM**, Heeren T, Beaver KA, Flynn M, **Battaglia TA**. Depression and anxiety diagnoses are not associated with delayed resolution of abnormal mammograms and pap tests among vulnerable women. J Gen Intern Med. 2012 Apr; 27(4):452-7. PMID: 22083552.

45. **Kronman AC**, **Paasche-Orlow M**, **Orlander JD**. Factors associated with disclosure of medical errors by housestaff. BMJ Qual Saf. 2012 Apr; 21(4):271-8. PMID: 21949438.

CHEST Supplement

ANTITHROMBOTIC THERAPY AND PREVENTION OF THROMBOSIS, 9TH ED: ACCP GUIDELINES

Oral Anticoagulant Therapy

Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines

Walter Ageno, MD; Alexander S. Gallus, MBBS; Ann Wittkowsky, PharmD, FCCP; Mark Crowther, MD; Elaine M. Hylek, MD, MPH; and Gualtiero Palareti, MD

Background: **The objective of this article is to summarize the published literature concerning the pharmacokinetics and pharmacodynamics of oral anticoagulant drugs that are currently available for clinical use and other aspects related to their management.**
Methods: **We carried out a standard review of published articles focusing on the laboratory and clinical characteristics of the vitamin K antagonists; the direct thrombin inhibitor, dabigatran etexilate; and the direct factor Xa inhibitor, rivaroxaban.**
Results: **The antithrombotic effect of each oral anticoagulant drug, the interactions, and the monitoring of anticoagulation intensity are described in detail and discussed without providing specific recommendations. Moreover, we describe and discuss the clinical applications and optimal dosages of oral anticoagulant therapies, practical issues related to their initiation and monitoring, adverse events such as bleeding and other potential side effects, and available strategies for reversal.**
Conclusions: **There is a large amount of evidence on laboratory and clinical characteristics of vitamin K antagonists. A growing body of evidence is becoming available on the first new oral anticoagulant drugs available for clinical use, dabigatran and rivaroxaban.**

CHEST 2012; 141(2)(Suppl):e44S–e88S

Abbreviations: AC = anticoagulation clinic; AMS = anticoagulation management service; aPTT = activated partial thromboplastin time; AUC = area under the curve; Cmax = peak plasma concentration; ECT = ecarin clotting time; HR = hazard ratio; INR = international normalized ratio; ISI = international sensitivity index; PCC = prothrombin complex concentrate; PE = pulmonary embolism; POC = point of care; PSM = patient self-management; PST = patient self testing; PT = prothrombin time; TCT = thrombin clotting time; TTR = time in therapeutic range; UC = usual care; VKA = vitamin K antagonist; VKOR = vitamin K oxide reductase; WHO = World Health Organization

For many decades, the vitamin K antagonists (VKAs) have been the only oral anticoagulant drugs available for clinical use for the primary and secondary prevention of venous and arterial thromboembolic events. VKAs have been consistently shown to be highly effective in many settings and are now used by millions of patients worldwide. Laboratory and clinical studies have contributed to understanding of the complex pharmacokinetics and pharmacodynamics of VKAs, their interactions, antithrombotic effects, and the risks associated with their use. Several studies have addressed the practical issues related to the management of patients on VKAs treatment, with particular focus on laboratory and clinical monitoring and on reversal strategies.

More recently, new oral anticoagulant drugs, namely the direct thrombin inhibitor dabigatran etexilate and the direct factor Xa inhibitor rivaroxaban, have been approved for clinical use in several countries. A growing body of laboratory and clinical data is becoming available to better understand the mechanisms of action and the optimal management of these new compounds. In this article we summarize the published literature concerning the pharmacokinetics and pharmacodynamics of all oral anticoagulant drugs that are currently available for clinical use and other aspects related to their management.

1.0 Vitamin K Antagonists

1.1 Pharmacology

VKAs produce their anticoagulant effect by interfering with the cyclic interconversion of vitamin K and its 2,3 epoxide (vitamin K epoxide), thereby modulating the γ-carboxylation of glutamate residues (Gla) on the N-terminal regions of vitamin K-dependent proteins (Fig 1).[1-8] The vitamin K-dependent coagulation factors II, VII, IX, and X require γ-carboxylation for their procoagulant activity, and treatment with VKAs results in the hepatic production of partially carboxylated and decarboxylated proteins with reduced coagulant activity.[9,10] Carboxylation is required for a calcium-dependent conformational change in coagulation proteins[11-13] that promotes binding to cofactors on phospholipid surfaces. In addition, the VKAs inhibit carboxylation of the regulatory anticoagulant proteins C, S, and Z and thereby have the potential to be procoagulant.[14] Although the anticoagulant effect of VKAs is dominant, a transient procoagulant effect may occur when baseline protein C and protein S levels are reduced due to the start of VKA therapy and the acute phase of a thrombotic event and before the balanced decrease of vitamin K-dependent clotting factor levels is achieved. Carboxylation requires the reduced form of vitamin K (vitamin KH_2), a γ-glutamyl carboxylase, molecular oxygen, and CO_2.[1] Vitamin K epoxide can be reused by reduction to VKH_2. The oxidation-reduction reaction involves a reductase pair. The first, vitamin K epoxide reductase, is sensitive to VKA, whereas vitamin K reductase is less sensitive.[1-3] Therefore, the anticoagulant effect of the VKAs can be overcome by low doses of phytonadione (vitamin K) (Fig 1).

1.2 Pharmacokinetics and Pharmacodynamics

Warfarin is a racemic mixture of two optically active isomers, the R and S enantiomers. Warfarin is highly water soluble, is rapidly absorbed from the gastrointestinal tract, has high bioavailability,[15,16] and reaches maximal blood concentrations about 90 min after oral administration.[15,17] Racemic warfarin has a half-life of 36 to 42 h[18] (R-warfarin 45 h, S-warfarin 29 h), circulates bound to plasma proteins (mainly albumin), and accumulates in the liver where the two enantiomers are metabolically transformed by different pathways (Fig 1).[18] The S enantiomer of warfarin (2.7-3.8 times more potent than the R enantiomer) undergoes approximately 90% oxidative metabolism, primarily by the CYP2C9 enzyme of the cytochrome P450 system and to a lesser extent by CYP3A4.[19] The less potent R enantiomer undergoes approximately 60% oxidative metabolism, primarily by two cytochrome P450 enzymes, CYP1A2 and CYP3A4, and to a lesser extent by CYP2C19. The remainder of the metabolism of both enantiomers involves reduction to diastereomeric alcohols. The relationship between the dose of warfarin and the response is modified by genetic and environmental factors that can influence the absorption of warfarin, its pharmacokinetics, and its pharmacodynamics.

Other available VKAs include acenocoumarol, phenprocoumon, and fluindione. Like warfarin, acenocoumarol and phenprocoumon also exist as optical isomers, but with different stereochemical characteristics. R-acenocoumarol has an elimination half-life of 9 h, is primarily metabolized by CYP2C9 and CYP2C19, and is more potent than S-acenocoumarol because of faster clearance of S-acenocoumarol, which has an elimination half-life of 0.5 h and is primarily metabolized by CYP2C9.[20] Phenprocoumon is a much longer-acting agent, with both the R- and S-isomers having elimination half-lives of 5.5 days. Both are metabolized by CYP2C9, and S-phenprocoumon is 1.5 to 2.5 times more potent than R-phenprocoumon.[21] Finally, fluindione is an indandione VKA with a mean half-life of 31 h.[22] Unlike warfarin, fluindione is not a chiral compound.[22]

1.3 Interactions

1.3.1 Genetic Factors: A number of point mutations in the gene coding for the CYP2C9 have been identified.[23] These polymorphisms, the most common of which are CYP2C9*2 and CYP2C9*3, are associated with an impaired ability to metabolize S-warfarin, resulting in a reduction in S-warfarin clearance and,

Revision accepted August 31, 2011.
Affiliations: From the University of Insubria (Dr Ageno), Varese, Italy; Flinders University (Dr Gallus), Adelaide, SA, Australia; the University of Washington (Dr Wittkowsky), Seattle, WA; McMaster University (Dr Crowther), St. Joseph's Hospital, Hamilton, ON, Canada; the Boston University School of Medicine (Dr Hylek), Boston, MA; and the University Hospital S. Orsola-Malpighi (Dr Palareti), Bologna, Italy.
Funding/Support: The Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines received support from the National Heart, Lung, and Blood Institute [R13 HL104758] and Bayer Schering Pharma AG. Support in the form of educational grants was also provided by Bristol-Myers Squibb; Pfizer, Inc; Canyon Pharmaceuticals; and sanofi-aventis US.

Correspondence to: Walter Ageno, MD, Department of Clinical Medicine, Ospedale di Circolo, Viale Borri 57, 21100 Varese, Italy; e-mail: walter.ageno@uninsubria.it

DOI: 10.1378/chest.11-2292

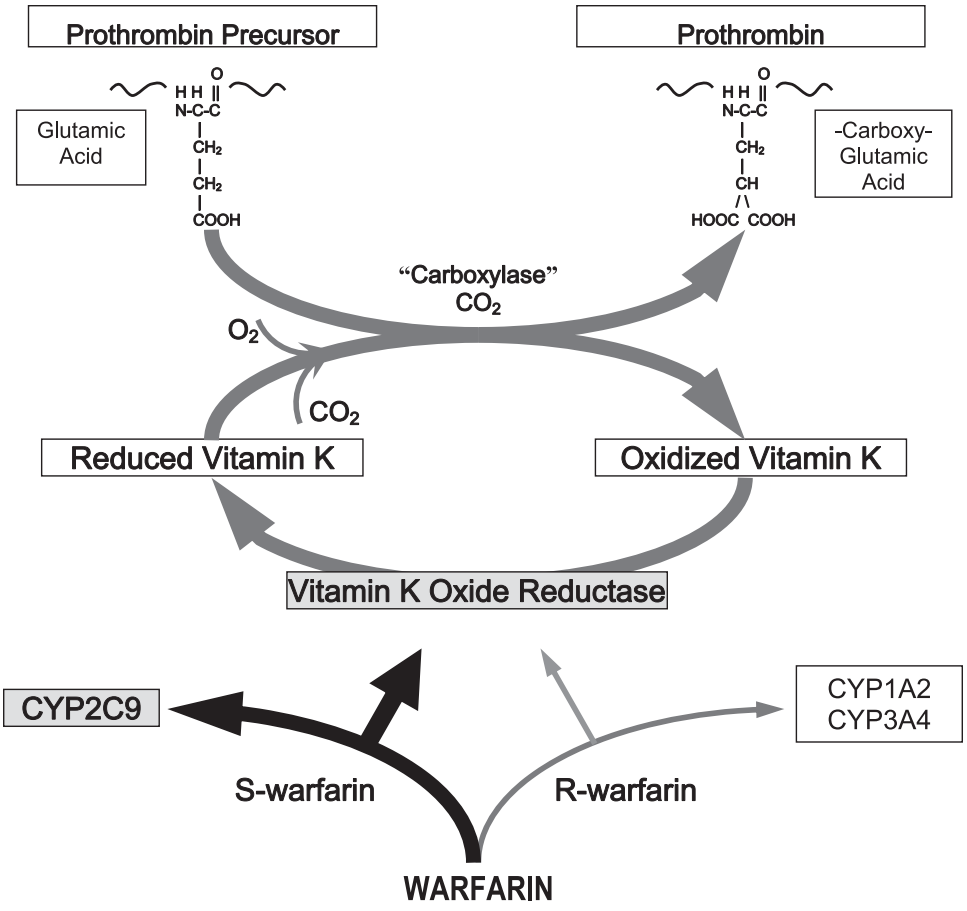


FIGURE 1. [Section 1.1] Vitamin K_1 is reduced to vitamin KH2. The major warfarin-sensitive enzyme in this reaction is the vitamin K oxide reductase mainly inhibited by the S-enantiomer of warfarin. S-warfarin is metabolized by the p450 cytochrome enzyme, CYP2C9. Reprinted with permission from Ansell et al.[8]

as a result, an increased S-warfarin elimination half-life.[24] Mutations in this gene occur with different frequencies in various ethnic groups (Table S1).[25,26] In comparison with patients who are homozygous for the wild-type allele (CYP2C9*1*1), patients with heterozygous (CYP2C9*1*2, CYP2C9*1*3, CYP2C9*2*3) or homozygous (CYP2C9*2*2, CYP2C9*3*3) expression of a variant allele require lower doses of warfarin, as determined by a systematic review of the literature and meta-analysis of studies that assessed the influence of CYP2C9 polymorphisms on warfarin dose requirements (Table S2).[27] Several investigations[25,28,29] have shown that these mutations, as well as others,[30-32] are also associated with an increase in bleeding complications associated with warfarin therapy. Mutations in CYP2C9 also affect acenocoumarol, although to a lesser degree because the anticoagulation potencies of the R and S enantiomers are comparable.[33,34] The effects of CYP2C9 polymorphisms are least pronounced with the use of phenprocoumon.[33,35]

The target for warfarin's inhibitory effect on the vitamin K cycle is the vitamin K oxide reductase (VKOR) enzyme first described in 1974.[36] The gene coding for the VKOR protein is located on the short arm of chromosome 16.[37,38] The gene encodes for several isoforms of a protein that are collectively termed the vitamin K oxide reductase complex 1 (VKORC1). Subsequently, mutations in this gene have been identified leading to enzymes with varying sensitivities to inhibition by warfarin,[38-43] thereby affecting the pharmacodynamics of warfarin. The mutations occur with differing frequencies in various ethnic populations and account, in part, for the difference in warfarin doses required to maintain a therapeutic international normalized ratio (INR) (Table S1) (tables that contain an "S" before the number denote supplementary tables not contained in the body of the article and available instead in an online data supplement; see the "Acknowledgments" for more information).[39-41,44,45]

Genetic mutations in the gene coding for the VKORC1often involve several mutations leading to various haplotypes that cause greater resistance to warfarin therapy. Harrington et al[43] found a warfarin-resistant individual who had high serum warfarin

concentrations and a 196G>A transition, predicting a Val66Met substitution in VKORC1. D'Andrea et al,[39] studying 147 patients, found that those with a 1173CC genotype required a higher mean maintenance dose compared with those with a CT or TT genotype, as did Quiteineh et al,[46] who found that a 1173 C>T polymorphism was significantly associated with the risk of anticoagulant overdose. By identifying a number of noncoding single nucleotide polymorphisms, Rieder et al[40] were able to infer that there are five major haplotypes associated with different dose requirements for maintaining a therapeutic INR. The maintenance dose ranged from a low of 2.7 mg warfarin per day for the sensitive haplotypes up to a high of 6.2 mg per day for the resistant haplotypes. Asian Americans had the highest proportion of sensitive haplotypes, whereas African Americans more frequently exhibited the resistant haplotypes (Table S1).

1.3.2 Drugs: VKAs are highly susceptible to drug-drug interactions. For warfarin, for example, manufacturer-provided product information lists >200 specific agents that may interfere with this agent.[47] Unfortunately, there seems to be little concordance among commonly used drug compendia and product labels with respect to interactions involving warfarin. Indeed, a major problem with the literature on this topic is that many reports are single-case reports and are not well documented. Anthony et al[44] recently reviewed three drug information compendia, Clinical Pharmacology, ePocrates, and Micromedex, and the warfarin sodium (Coumadin) product label approved by the US Food and Drug Administration, for listings of interactions between warfarin and drugs, biologics, foods, and dietary supplements and found that of a total of 648 entries from the four sources, only 50 were common to all the sources.[44] As in the previous edition of this article,[8] Table 1 summarizes a comprehensive list of drugs that potentiate, inhibit, or have no effect on the anticoagulant effect of warfarin based on the results of a systematic review of available evidence completed in 2005, which rated warfarin drug interaction reports according to interaction direction, clinical severity, and quality of evidence, and developed lists of warfarin drug interactions considered highly probable, probable, possible, and highly improbable.[48]

Drugs such as cholestyramine can reduce the anticoagulant effect of warfarin by reducing its absorption. Other drugs potentiate the anticoagulant effect of warfarin by inhibiting its clearance, whereas some drugs may inhibit the anticoagulant effect by enhancing its clearance.[49] These latter effects may be through stereoselective or nonselective pathways,[50,51] (stereoselective interactions may affect the oxidative metabolism of either the S-enantiomer or R-enantiomer of warfarin). The inhibition of S-warfarin metabolism is more important clinically, because this enantiomer is more potent than the R-enantiomer as a VKA.[50,51] Phenylbutazone,[52] sulfinpyrazone,[53] metronidazole,[54] and trimethoprimsulfamethoxazole[55] inhibit the clearance of S-warfarin, and each potentiates the effect of warfarin on the prothrombin time (PT). In contrast, drugs such as cimetidine and omeprazole, which inhibit the clearance of the R-isomer, potentiate the PT only modestly in patients who are treated with warfarin.[51,54,56] Amiodarone is a potent inhibitor of the metabolic clearance of both the S-enantiomer and the R-enantiomer and potentiates warfarin anticoagulation.[57] The anticoagulant effect of warfarin is inhibited by drugs like barbiturates, rifampin, azathioprine, and carbamazepine, which increase its clearance by inducing hepatic metabolism.[58] Azathioprine also reduces the anticoagulant effect of warfarin, presumably through a potentiating effect on hepatic clearance.[59] Long-term alcohol consumption has a similar potential to increase the clearance of warfarin, but ingestion of even relatively large amounts of wine had little influence on the PT in normal volunteers who were given warfarin.[60] The effect of enzyme induction on warfarin therapy has been analyzed in a critical review.[58] Ten hepatic microsomal enzyme agents were assessed. Enzyme induction of warfarin metabolism by rifampin and barbiturates was considered likely, and an interaction with carbamazepine, griseofulvin, aminoglutethimide, nafcillin, and dicloxacillin was considered probable.

Drugs may also influence the pharmacodynamics of warfarin by inhibiting the synthesis of or increasing the clearance of vitamin K-dependent coagulation factors or by interfering with other pathways of hemostasis. The anticoagulant effect of warfarin is augmented by second-generation and third-generation cephalosporins, which inhibit the cyclic interconversion of vitamin K;[61,62] by thyroxine, which increases the metabolism of coagulation factors;[63] and by clofibrate through an unknown mechanism.[64] Doses of salicylates of >1.5 g per day[65] may augment the anticoagulant effect of warfarin. Acetaminophen potentiates the effect of warfarin when used over prolonged periods of time, as demonstrated in a recent randomized, blinded trial.[66-68] Acetaminophen possibly potentiates the anticoagulant effect of warfarin through inhibition of VKOR by a toxic metabolite of the drug,[69] although the accumulation of this metabolite may vary among individuals, thus accounting for a variable potentiating effect.[70] Heparin potentiates the anticoagulant effect of warfarin, but in therapeutic doses produces only a slight prolongation of the PT. The mechanisms by which erythromycin[71] and some anabolic steroids[72] potentiate the anticoagulant effect of

Table 1—*[Section 1.3.2] Drug, Food, and Dietary Supplement Interactions With Warfarin by Level of Supporting Evidence and Direction of Interaction*

Level of Causation	Anti-infectives	Cardiovascular	Analgesics, Antiinflammatories, and Immunologics	CNS Drugs	GI Drugs and Food	Herbal Supplements	Other Drugs
			Potentiation				
Highly probable	Ciprofloxacin Cotrimoxazole Erythromycin Fluconazole Isoniazid Metronidazole Miconazole oral gel Miconazole vaginal suppository Voriconazole	Amiodarone Clofibrate Diltiazem Fenofibrate Propafenone Propranolol Sulfinpyrazone (biphasic with later inhibition)	Phenylbutazone Piroxicam	Alcohol (if concomitant liver disease) Citalopram Entacapone Sertraline	Cimetidine Fish oil Mango Omeprazole	Boldo-fenugreek Quilinggao	Anabolic steroids Zileuton
Probable	Amoxicillin/clavulanate Azithromycin Clarithromycin Itraconazole Levofloxacin Ritonavir Tetracycline	Aspirin Fluvastatin Quinidine Ropinirole Simvastatin	Acetaminophen Aspirin Celecoxib Dextropropoxyphene Interferon Tramadol	Disulfiram Chloral hydrate Fluvoxamine Phenytoin (biphasic with later inhibition)	Grapefruit	Danshen Don quai Lycium barbarum L PC-SPES	Fluorouracil Gemcitabine Levamisole/fluorouracil Paclitaxel Tamoxifen Tolterodine
Possible	Amoxicillin Amoxicillin/tranexamic rinse Chloramphenicol Gatifloxacin Miconazole topical gel Nalidixic acid Norfloxacin Ofloxacin Saquinavir Terbinafine	Amiodarone-induced toxicosis Disopyramide Gemfibrozil Metolazone	Celecoxib Indomethacin Leflunomide Propoxyphene Rofecoxib Sulindac Tolmetin Topical salicylates	Felbamate	Orlistat	Danshen/methyl salicylates	Acarbose Cyclophosphamide/methotrexate/fluorouracil Daptomycin Danazol Iphosphamide Trastuzumab
Highly improbable	Cefamandole Cefazolin Sulfisoxazole	Bezafibrate Heparin	Levamisole Methylprednisolone Nabumetone	Fluoxetine/diazepam Quetiapine			Etoposide/carboplatin Levonorgestrel
			Inhibition				
Highly probable	Griseofulvin Nafcillin Ribavirin Rifampin	Cholestyramine	Mesalamine	Barbiturates Carbamazepine	High vitamin K content foods/enteral feeds Avocado (large amounts)		Mercaptopurine

(Continued)

Table 1—*Continued*

Level of Causation	Anti-infectives	Cardiovascular	Analgesics, Antiinflammatories, and Immunologics	CNS Drugs	GI Drugs and Food	Herbal Supplements	Other Drugs
Probable	Dicloxacillin Ritonavir	Bosentan	Azathioprine	Chlordiazepoxide	Soy milk Sucralfate	Ginseng	Chelation therapy Influenza vaccine Multivitamin supplement Raloxifene HCL
Possible	Terbinafine	Telmisartan	Sulfasalazine		Sushi containing seaweed		Cyclosporine Etretinate Ubidecarenone
Highly improbable	Cloxacillin Nafcillin/dicloxacillin Teicoplanin	Furosemide		Propofol		Green tea	

Data from Holbrook et al.[48]

warfarin are unknown. Sulfonamides and several broad-spectrum antibiotic compounds may augment the anticoagulant effect of warfarin in patients consuming diets that are deficient in vitamin K by eliminating bacterial flora and aggravating vitamin K deficiency.[73]

Drugs such as aspirin,[74] nonsteroidal antiinflammatory drugs,[75,76] penicillins in high doses,[77,78] and moxalactam[62] increase the risk of warfarin-associated bleeding by inhibiting platelet function. Of these, aspirin is the most important because of its widespread use and prolonged effect.[79,80] Aspirin and nonsteroidal antiinflammatory drugs can also produce gastric erosions that increase the risk of upper GI bleeding. This can occur even with COX-2 inhibitors, which were originally believed to be less likely to predispose to gastric bleeding than nonsteroidal antiinflammatory drugs.[76] In one case-controlled analysis of 98,821 subjects on warfarin identified in linked databases, celecoxib and rofecoxib were associated with a 1.7- or 2.4-fold increased risk of GI hemorrhage, respectively.[76] The risk of clinically important bleeding is heightened when high doses of aspirin are taken during high-intensity warfarin therapy (INR, 3.0-4.5).[74,81] However, low doses of aspirin (ie, 75 to 100 mg daily) combined with moderate-intensity and low-intensity warfarin anticoagulation therapy are also associated with increased rates of bleeding.[82,83] The effect of statins or fibrates on the risk of bleeding in patients on VKAs is controversial. The initiation of a fibrate or statin that inhibits CYP3A4 enzymes was recently reported to increase the risk of gastrointestinal bleeding in warfarin-treated patients, whereas other statins that are mainly excreted unchanged were not found to be associated with such an increased risk.[84] Conversely, other authors reported that long-term statin use is associated with a decreased risk of bleeding complications in patients with atrial fibrillation (AF) on warfarin therapy.[85]

The most effective method to avoid adverse outcomes associated with drug interactions is to try to avoid, when feasible, concurrent use of potentially interacting drugs and to use noninteracting alternatives instead. When noninteracting alternatives are not available, adverse outcomes can be avoided by increasing the frequency of monitoring and adjusting warfarin doses based on INR response. Prospective dosing adjustments are inappropriate because of the unpredictable nature of patient response to drug interactions (see Holbrook et al[86] in this supplement).

1.3.3 Environmental Factors: Nutritional supplements and herbal products are particularly problematic in that warfarin-treated patients often fail to inform physicians that they are using such products and physicians rarely ask. In one survey of 1,200 patients from four large anticoagulation clinics (ACs) in the

United States, one-third of the patients used dietary supplements and one-third of all patients surveyed indicated that their provider failed to discuss potential interactions with them.[87] There is also little or no standardization of the content of such products, especially herbal remedies, and reports of interactions are often anecdotal or single-case reports, without good substantiation.[88-91] Of the higher-quality studies, ginkgo and ginger were shown not to have an effect on the pharmacokinetics and pharmacodynamics of warfarin in healthy subjects in a randomized, open-label, crossover, study,[92] and coenzyme Q_{10} (and ginkgo) was shown not to have an effect on warfarin dosage in a randomized, double-blind, crossover study.[93] Ginseng was shown to reduce the effect of warfarin in a randomized, placebo-controlled trial.[94] Not surprisingly, products such as green tea, with a high content of vitamin K, were shown to reduce the anticoagulant effect of warfarin.[48]

Subjects receiving long-term warfarin therapy are sensitive to fluctuating levels of dietary vitamin K,[95,96] which is derived predominantly from phylloquinones in plant material.[96] Sadowski and associates[97] have listed the phylloquinone content of a wide range of food, and the list can be found on the Internet (http://ods.od.nih.gov/factsheets/cc/coumadin1.pdf). Phylloquinones act through the warfarin-insensitive pathway.[98] Important fluctuations in vitamin K intake can occur in both healthy and sick subjects.[99] An increased intake of dietary vitamin K that is sufficient to reduce the anticoagulant response to warfarin occurs in patients consuming green vegetables or vitamin K-containing supplements, during weight-reduction diets, and in patients who have been treated with vitamin K supplements.[100] Reduced dietary vitamin K intake potentiates the effect of warfarin in ill patients who have been treated with antibiotics and IV fluids without vitamin K supplementation and in patients who have states of fat malabsorption.

In general, a consistent intake of vitamin K-containing foods is advisable, but neither specific restrictions nor additions seem necessary in patients with stable anticoagulant control. Patients should be informed of possible changes in INR, in particular in response to the use of dietary supplements or herbs, or alcohol used chronically or ingested in large quantities.[101,102] More frequent monitoring of the INR should be proposed if dietary habits have substantially changed in response to weight reduction diets, periods following hospitalization, treatment with chemotherapy, sustained diarrhea or vomiting, or in case of anorexia.[103]

A number of other conditions and disease states have been observed to influence anticoagulation with warfarin. Hepatic dysfunction potentiates the response to warfarin through the impaired synthesis of coagulation factors.[104] These patients may appear to be "auto-anticoagulated" with baseline elevated INRs, but the degree of suppression of clotting factors does not mimic that of patients treated with warfarin and is not sufficient to prevent thromboembolism.[105] Hypermetabolic states produced by fever or hyperthyroidism increase warfarin responsiveness, probably by increasing the catabolism of vitamin K-dependent coagulation factors.[70,106] Certain components of cigarette smoke may induce CYP1A2, and warfarin dosing requirements have been observed to decline after smoking cessation.[107,108] Chewing tobacco contains high quantities of vitamin K that can increase warfarin dose requirements.[109] Exacerbations of heart failure can increase responsiveness to warfarin therapy, likely in response to the effect of hepatic congestion on warfarin metabolism.[110] End-stage renal disease is associated with reduced activity of CYP2C9, leading to lower warfarin dosing requirements in these patients.[111] Warfarin dosing requirements decline with advanced age as a result of reduced availability of vitamin K stores and lower plasma concentrations of vitamin K-dependent clotting factors.[112-114] In fact, age may be the single most important easily obtained predictor of warfarin dosing requirement.[115]

1.4 Antithrombotic Effect

The antithrombotic effect of VKAs is attributed to their anticoagulant effect, which in turn is mediated by the reduction of four vitamin K-dependent coagulation factors. The experiments of Wessler and Gitel[116] >40 years ago using a stasis model of thrombosis in rabbits showed that the antithrombotic effect of warfarin requires 6 days of treatment and requires the reduction of prothrombin (factor II), which has a relatively long half-life of about 60 to 72 h, compared with 6 to 24 h for other vitamin K-dependent factors. In a rabbit model of tissue factor-induced intravascular coagulation, the protective effect of warfarin mainly resulted from lowering prothrombin levels.[117] Patel and associates[118] demonstrated that clots formed from umbilical cord plasma containing about half the prothrombin concentration of plasma from adult control subjects generated significantly less fibrinopeptide A than clots formed from maternal plasma. The view that warfarin exerts its antithrombotic effect by reducing prothrombin levels is consistent with observations that clot-bound thrombin is an important mediator of clot growth[119] and that reduction in prothrombin levels decreases the amount of thrombin generated and bound to fibrin, thereby reducing thrombogenicity.[118]

The suggestion that the antithrombotic effect of VKAs is reflected in lower levels of prothrombin forms the basis for overlapping the administration of

a parenteral anticoagulant with warfarin until the PT or INR is prolonged into the therapeutic range during the treatment of patients with thrombosis. Since the half-life of prothrombin is about 60 to 72 h, at least 5 days of overlap is necessary.

1.5 Monitoring Anticoagulant Intensity: the INR

The PT test[120] is the most common test used to monitor VKA therapy. The PT responds to a reduction of three of the four vitamin K-dependent procoagulant clotting factors (ie, II, VII, and X) that are reduced by warfarin at a rate proportional to their respective half-lives. Thus during the first few days of warfarin therapy the PT reflects mainly a reduction of factor VII, the half-life of which is approximately 6 h. Subsequently, the reduction of factors X and II contributes to prolongation of the PT. The PT assay is performed by adding calcium and thromboplastin to citrated plasma. Thromboplastins vary in responsiveness to a reduction of the vitamin K-dependent coagulation factors. An unresponsive thromboplastin produces less prolongation of the PT for a given reduction in vitamin K-dependent clotting factors than a responsive one. The responsiveness of a thromboplastin can be measured by assessing its international sensitivity index (ISI) (see later discussion in this section). Highly sensitive thromboplastins (indicated by an ISI of approximately 1.0) are now available that are composed of human tissue factor produced by recombinant technology and defined phospholipid preparations.

PT monitoring of VKA treatment is not standardized when expressed in seconds, or as a simple ratio of the patient plasma value to that of plasma from a healthy control subject, or as a percentage of diluted normal plasma. A calibration model,[121] which was adopted in 1982, is now used to standardize reporting by converting the PT ratio measured with the local thromboplastin into an INR, calculated as follows:

$$\text{INR} = (\text{patient PT}/\text{mean normal PT})^{\text{ISI}}$$

or

$$\log \text{INR} = \text{ISI}(\log \text{observed PT ratio})$$

where ISI denotes the ISI of the thromboplastin used at the local laboratory to perform the PT measurement. The ISI reflects the responsiveness of a given thromboplastin to the reduction of the vitamin K-dependent coagulation factors compared with the primary World Health Organization (WHO) international reference preparations, so that the more responsive the reagent, the lower the ISI value.[121] It is proposed that patients' samples for calibration should be selected by rejecting samples beyond the 1.5 to 4.5 INR range.[122] Outliers, defined as points with a perpendicular distance greater than three residual SDs from the line of relationship, should be rejected. Selection of patients' samples and rejection of outliers results in a reduction of the between-laboratory variation of calibration.[122]

As the INR standard of reporting was widely adopted, a number of problems surfaced. These are listed in Table 2 and are reviewed briefly here.

The INR is based on ISI values derived from the plasma of patients who had received stable anticoagulant doses for at least 6 weeks.[123] As a result, the INR has not been validated and should be viewed with some skepticism early in the course of warfarin therapy, particularly when results are obtained from different laboratories. Even under these conditions,

Table 2—*[Section 1.5] Potential Problems With the INR (Causes of Erroneous INR)*[8]

Problems	Description
1. Incorrect PTR from erroneous PT determination due to	
Pretest variables (sampling and blood collection problems)	Trisodium-citrate concentration, storage time, storage temperature, evacuated tube effects, inadequate sample, variations in manual technique
Incorrect normal value	From nonuse of MNPT, error in MNPT due to unrepresentative selection, technical faults (see above), nonuse of geometric mean
2. Incorrect ISI of local thromboplastin reagent/test system from lack of reliability of the ISI result provided by the manufacturer	Incorrect choice of IRP, poor distribution of coumarin test samples across treatment range, inadequate numbers of test samples in ISI calibration, incorrect transformation of PTR of test plasmas to INR
3. Drift of ISI since original calibration	
4. Instrument (coagulometer) effects on INR at local site	
5. Lupus anticoagulant effects on some thromboplastin reagents	
6. Lack of reliability of the INR system when used at the onset of warfarin therapy and for screening for a coagulopathy in patients with liver disease	
7. Relative lack of reliability of INR > 4.5 as these values excluded from ISI calibrations	

INR = international normalized ratio; IRP = international reference preparation; ISI = international sensitivity index; MNPT = mean normal prothrombin time; PT = prothrombin time; PTR = prothrombin time ratio.

however, the INR is more reliable than the unconverted PT ratio,[124] and its use is thus recommended during both the initiation and maintenance of VKAs.

The validity of the INR in other conditions of impaired coagulation has been less frequently evaluated. Some authors have recently challenged the use of the INR in patients with liver disease and in particular in prognostic scores such as the Model for End-stage Liver Disease.[125,126] Thus, a new INR specific for liver diseases has been proposed, derived by using plasma from patients with liver diseases to calibrate thromboplastins instead of plasma from patients on oral anticoagulant treatment with VKAs.[125,126]

The accuracy of the INR can be influenced by reagents with different sensitivities[127] and also by the automated clot detectors now used in most laboratories.[128-135] In general, the College of American Pathologists has recommended[136] that laboratories should use thromboplastin reagents that are at least moderately responsive (ie, ISI < 1.7) and reagent/instrument combinations for which the ISI has been established and validated.

ISI values provided by the manufacturers of thromboplastin reagents are not invariably correct when applied locally,[137-139] and this adversely affects the reliability of measurements. Local calibrations can be performed using plasma samples with certified PT values to determine the instrument-specific ISI. The mean normal plasma PT is not interchangeable with a laboratory control PT.[140] Therefore, the use of other than a properly defined mean normal PT can yield erroneous INR calculations, particularly when less-responsive reagents are used. The mean normal PT should be determined for each new batch of thromboplastin with the same instrument used to assay the PT.[140]

The concentration of citrate that is used to anticoagulate plasma may affect the INR.[141,142] In general, higher citrate concentrations (eg, 3.8%) lead to higher INR values,[141] and underfilling the blood collection tube spuriously prolongs the PT because excess citrate is present. Using collection tubes containing 3.2% concentrations of citrate for blood coagulation studies and adequately filling tubes can reduce this problem.

1.6 Practical Issues Related to Initiation and Maintenance

1.6.1 Approaches to the Induction of VKAs: Following the administration of VKAs, an initial effect on the INR usually occurs within the first 2 or 3 days, depending on the dose administered, and an antithrombotic effect occurs within the next several days.[143,144] There is room for flexibility in selecting a starting dose of warfarin. The results of clinical studies suggest that initiation doses between 5 and 10 mg are effective,[143,145,146] with individual responses varying according to the inpatient or outpatient status, age, concomitant treatments, and comorbidities. Thus, starting doses of ≤ 5 mg might be appropriate in the elderly, in patients with impaired nutrition, liver disease, or congestive heart failure, and in patients who are at high risk of bleeding.[112,115,147,148] An initial dose of 2 to 3 mg seems to be appropriate for patients who have undergone heart valve replacement, given their higher sensitivity to VKAs probably caused by the effects of cardiopulmonary bypass and concomitant therapies.[149] Recommendations are provided in Holbrook et al.[86]

As described in section 1.2.1, CYP2C9 genotype and VKORC1 haplotype influence warfarin dosing requirements. In response to these observations, numerous investigators have attempted to assess the combined influence of patient factors on warfarin dosing requirement, with a goal of developing models to predict warfarin dose requirements based on assessment of various environmental and genetic factors.

Sconce et al[42] found that a combination of CYP2C9 and VKORC1 genotypes plus height produced the best predictive model for estimating warfarin dose, whereas Vecsler et al[150] reported that CYP2C9 and VKORC1 genotypes together with age and body weight could explain as much as 63% of the dose variance, and Herman et al[151] could attribute 60% of dose variability to CYP2C9 and VKORC1polymorphisms, age, and body surface area. Limdi et al[152] found that CYP2C9 and VKORC1 polymorphisms accounted for 30% of the variability in warfarin dose among European Americans but only for 10% among African Americans. In a prospective cohort study of 48 consecutive patients starting warfarin for postorthopedic surgery prophylaxis, the patients with a variant CYP2C9 allele had a greater than fourfold increase in the risk of the INR exceeding 4.0. However, this substantial increase in the risk of supratherapeutic INR was observed despite the fact that all patients had their warfarin dose selected using a complex algorithm that considered CYP2C9 genotype.[153]

Gage et al[154] developed a dosing algorithm based on CYP2C9 and VKORC1 polymorphisms along with clinical and demographic factors. In the derivation cohort of 1,015 patients on warfarin therapy, body surface area, age, target INR, amiodarone use, smoker status, race, current thrombosis, VKORC1 polymorphism 1639/3673 G>A, CYP2C9(*)3, and CYP2C9(*)2 were all independent predictors of warfarin therapeutic dose.

Three prospective studies have compared the time in therapeutic range (TTR; see section 1.6.2 "Evaluating the Quality of Monitoring: TTR") of patients using a pharmacogenetics-based dosing strategy with the TTR achieved when patients were managed

without knowledge of genotype.[155-157] Two of these studies showed no difference in TTR, whereas one study reported increased times in range for the pharmacogenetic group but had significant design flaws. More recently, Klein et al[158] developed two dosing algorithms: one based on clinical variables only and one based on both clinical variables and genetic information. The ability of these two algorithms to predict subsequent warfarin doses was then compared in a validation cohort of > 1,000 patients. The pharmacogenetic algorithm more accurately identified patients who required ≤ 21 mg weekly warfarin doses and patients who required weekly doses of ≥ 49 mg in comparison with the clinical algorithm and to a fixed-dose approach, whereas no difference was detected in the prediction of intermediate doses.

A few studies have suggested that certain genotypes are associated with adverse events. Thus, Higashi et al[29] studied 185 patients, 58 with at least one variant genotype of CYP2C9, and found an increased risk of having INRs above range (hazard ratio [HR], 1.40; 95% CI, 1.03-1.90) and of a serious or life-threatening bleeding event (HR, 2.39; 95% CI, 1.18-4.86) in those with variant genotypes. The latter hazard estimate was based on a few events in a very small number of patients with the variant genotypes. Joffe et al,[159] also studying CYP2C9 single nucleotide polymorphisms, found a trend toward increased rates of an INR > 6.0 and of bleeding in patients who were categorized as heterozygotes, or compound heterozygotes/homozygotes, compared with those categorized as wild type, as did Veenstra et al.[30] A similar increased risk of bleeding was seen in patients with these polymorphisms who were taking acenocoumarol but not phenprocoumon.[32] On the other hand, neither CYP2C9 nor VKORC1influenced the risk of bleeding in a more recent study by Limdi et al.[160]

Likewise, the only high-quality, randomized, controlled trial performed to date showed that use of a pharmacogenetic-based dosing strategy did not significantly reduce the risk of adverse events (34.7% in pharmacogenetic group vs 42.4% in control group; OR = 0.72; 95% CI, 0.41-1.28]).[155] This issue is further discussed and recommendations are provided in Holbrook et al.[86]

When rapid anticoagulant effect is required, a rapidly acting parenteral anticoagulant should be started together with the VKA and discontinued after at least 5 days of concomitant therapy and once the INR has been in the therapeutic range for at least two measurements approximately 24 h apart. This allows factors X and II to be reduced to levels sufficient to treat VTE. If there is no urgent need for an immediate anticoagulant effect (eg, in chronic stable AF), warfarin administration can be commenced without the concurrent use of a rapid-acting anticoagulant.

1.6.2 Evaluating the Quality of Monitoring: TTR: The relationship between the intensity of treatment and the risk of an adverse event has been evaluated by examining the frequency of such events as a function of the TTR.[161-163] A strong relationship between TTR and the rates of bleeding or thromboembolic events has been observed across studies[161-173] with different patient populations, different target ranges, different scales for measuring intensity of anticoagulation (ie, PT, PT ratio, and INR), different methods of measuring TTR, and different models of dose management. In a large, retrospective analysis of patients with mechanical heart valves, Cannegieter et al[164] found that risks of major bleeding or thromboembolism were greatly increased during the times when patients were above or below the therapeutic range of INR compared with times when they were within range. A similar relationship has been demonstrated for other groups of patients.[169,174] A recent substudy examined the influence of TTR on the relative effectiveness of warfarin and dual antiplatelet therapy with aspirin plus clopidogrel in patients with nonvalvular AF, in which the overall result favored warfarin.[165] In this large retrospective analysis, the TTR during warfarin therapy appeared to be a major determinant of its efficacy, since the advantage of warfarin over antiplatelet therapy was lost below a threshold TTR of between 58% and 65%.[165] The percentage of INRs or TTR is highly dependent on the quality of dose management as reflected in studies that report TTR. Poor quality of dose management results in a high proportion of low INRs during the first 3 months of treatment following an acute DVT, which in turn predicts a higher rate of subsequent recurrence.[161,175] The TTR reflects the quality of dose adjustment in studies of patients managed in a usual care (UC) setting, by an anticoagulation management service (AMS), by patient self testing (PST) or patient self management (PSM), or in the setting of a randomized trial.

TTR can be determined in different ways, so comparisons between studies may be difficult.[176] TTR is most commonly estimated by using one of three methodologies: calculating the fraction of all INR values that are within the therapeutic range (ie, the number of INRs in range divided by the total number of INR tests); using the "cross-section of the files" methodology, which assesses the fraction of patients with an INR in range at one point in time compared with the total number of patients who had an INR measured at that point in time; or applying the linear interpolation method of Rosendaal et al,[177] which assumes that a linear relationship exists between two INR values and allocates a specific INR value to each day between tests for each patient. Each approach has its advantages and disadvantages.[176] Furthermore, the results of all these methods depend on whether

an exact or an expanded therapeutic range is used,[178] whether INRs obtained during invasive procedures when warfarin therapy might be interrupted are included, and whether different oral anticoagulant preparations (eg, warfarin, phenprocoumon, acenocoumarol, or fluindione) are included.[179,180] Since clinical outcome studies have not compared one methodology with another and correlated their results with adverse events, no one method can be recommended, and the reader should be aware of these differences.

1.6.3 Frequency of Monitoring: In hospitalized patients, INR monitoring is usually performed daily until the therapeutic range has been achieved and maintained for at least 2 consecutive days. In outpatients starting VKA therapy, initial monitoring may be reduced to once every few days until a stable dose response has been achieved. When the INR response is stable, the frequency of testing can be reduced to intervals as long as every 4 to 6 weeks (or possibly longer in particularly stable patients). If adjustments to the dose are required, then the cycle of more frequent monitoring should be repeated until a stable dose response can again be achieved.

The optimal frequency of long-term INR monitoring is influenced by patient compliance, transient fluctuations in the severity of comorbid conditions, the addition or discontinuation of other medications, changes in diet, the quality of dose-adjustment decisions, and whether the patient has demonstrated a stable dose response. The dose required to maintain a therapeutic range for patients >60 years of age decreases with increasing age,[112,115,181,182] possibly because of a reduction in the clearance of warfarin with age.[183] Gender also influences dose, with women requiring less warfarin to maintain a therapeutic INR than men at an equivalent age.[115]

To compare different intervals between measurements of INR in stable patients, Pengo et al[184] randomized 124 patients with prosthetic mechanical heart valves and at least 6 months of a steady dose requirement to INR monitoring at either 6-week or 4-week intervals. They found no differences of time in, above, or below range between the groups, although the actual monitoring intervals were 24.9 days in the 6-week group and 22.5 days in the 4-week group ($P<.0003$).[184] In a retrospective study of $>4{,}000$ patients with chronic AF and $>250{,}000$ INRs, Shalev et al[185] found that time in range increased as the testing interval decreased from every 5 weeks or more to every 3 weeks (41% to 48%, $P<.0005$), and the investigators suggested that patients should be monitored at time intervals no longer than every 3 weeks. However, the strength of the recommendation is reduced by the very poor TTR reported in both arms of the study. More recently, Witt et al[186,187] found that patients with very stable INR levels defined by extremely high TTR rates required fewer visits for INR control and had significantly fewer hemorrhagic and thromboembolic events than patients with a less-stable INR; they suggested that many warfarin-treated patients whose INR values remain within the therapeutic range over time could be safely treated with INR recall intervals >4 weeks. In this study, advanced age predicted stable anticoagulation.

1.6.4 Factors Associated With INR Stability in Long-term Management: Two recent studies have assessed factors associated with very stable INR control during treatment with VKAs.[186,187] In the first study, 2,504 patients with INR values entirely within the INR range for 6 months were compared with 3,569 patients with at least one INR value outside the INR range.[186] In the second study, 533 patients with INR values within the therapeutic range for 12 months were compared with 2,555 control subjects.[187] Independent predictors of stability were age >70 years, the absence of chronic diseases, and (in one study only[187]) male gender. Congestive heart failure, diabetes, and a target range for INR ≥ 3.0 were associated with instability.

Physical activity also seems to play a role in the stability of the response to warfarin. A reduction in the anticoagulant effect has been found to be correlated with a sudden increase in physical activity. An increase in warfarin requirements associated with an increase in physical activity (represented by a daily exercise such as walking) has been described both in patients and in healthy subjects.[188,189]

Changes in dietary vitamin K intake may influence the stability of the INR in patients on VKAs, and a few trials have assessed its impact on therapeutic stability. Sorano et al[190] showed that controlling the intake of dietary vitamin K can achieve a more stable anticoagulant response. Sconce et al,[191] by comparing the daily vitamin K intake in 26 unstable patients and in 26 stable control patients, showed that unstable patients have poorer intake of vitamin K. Kurnik et al[192] showed that in vitamin K-depleted patients, very small amounts of vitamin K-containing vitamins will influence the INR to a greater extent compared with those with an adequate vitamin K status. Schurgers et al,[193] studying healthy volunteers on oral anticoagulation, found that a daily dose of vitamin K of at least 150 μg was needed to alter the INR response. Reese et al,[194] in a retrospective analysis, assessed the effect of a daily dose of 100 μg of vitamin K_1 in nine unstable patients. These patients experienced an increase in the percentage of INRs in range from 32% to 57% in response to the daily vitamin K. In a prospective, open-label, crossover study, Ford et al[195] found that

five of nine patients improved their stability with administration of low-dose vitamin K. As expected, the INR initially decreased in patients given vitamin K, and an increased dose of warfarin was needed to reestablish an INR within the therapeutic range, which took from 2 to 35 days to achieve. In a nested case-control study, Rombouts et al[196] assessed the effect of dietary vitamin K intake on the risk of subtherapeutic INR values and the interaction between usual and recent vitamin K intake. Patients with a high usual vitamin K intake had less risk of a subtherapeutic INR, an effect possibly mediated by a reduced influence on the INR of an incidental consumption of vitamin K-rich food when there is a high usual intake of vitamin K. In these last two studies, variable INR was defined as requiring a minimum of three warfarin dose changes or three INRs outside of the therapeutic range in the preceding 6 months,[195] or an INR SD > 0.5 with at least three warfarin dose changes during the previous 6 months.[196]

Sconce et al[197] conducted the first randomized, blinded trial in 70 unstable patients over a 6-month period. Vitamin K supplementation with 150 μg/d resulted in a significantly greater decrease in SD of the INR compared with placebo (-0.24 ± 0.14 vs -0.11 ± 0.18; $P < .001$) and a significantly greater increase in percentage of time within target INR range (28% ± 20% vs 15% ± 20%; $P < .01$). Finally, Rombouts et al[198] randomized 100 patients treated with phenprocoumon to receive daily doses of 100 μg of vitamin K and 100 patients to receive a placebo. Vitamin K improved the stability of anticoagulant therapy, with a relative risk of maximal stability (where all INR results were in range) in the vitamin K group compared with the placebo group of 1.8 (95% CI, 1.1-2.7).

De Assis et al[199] randomized patients with a recent INR value outside the therapeutic range to a conventional approach based on changes in anticoagulant prescription or to a strategy that adjusted the dietary intake of vitamin K. Patients in the latter group had significantly more INR values within the therapeutic range and significantly fewer episodes of minor bleeding as compared with patients in the former group.

As with any other drug, a patient's nonadherence to prescribed dosing with VKAs is one of the most important causes of INR instability. In a prospective cohort study at three US ACs aimed to determine the effect of adherence on anticoagulation control, patients treated with warfarin were monitored with an electronic system that records each time they open their medication container.[200] There was a high proportion of missed tablets and, as expected, in multivariable analyses there was a significant association between under-adherence and under-anticoagulation. In a case-control study on the most unstable patients from 35 Italian ACs, poor comprehension of the indications and mechanisms of VKAs was the most important predictor of instability among other factors, including working status, type of VKA, and a poor score on a mental test.[161] In a single-center study carried out on outpatients attending an AC in the United States, perceived barriers to compliance with warfarin, marital status, living arrangements, and drug regimen played significant roles in warfarin noncompliance.[201] In a prospective cohort study of adults initiating warfarin at two ACs, independent risk factors for nonadherence were education level, employment status, mental health functioning, and cognitive impairment.[202]

1.7 Data Management

An obstacle to the safety and effectiveness of warfarin therapy is the poor quality of dose management in routine clinical practice. Adequate anticoagulant care with the VKAs requires a system of patient education and careful data management to record and track INR values and to ensure patients are treated with anticoagulants for an appropriate period of time.

1.7.1 The Role of Anticoagulation Clinics: Nonrandomized, retrospective studies have reported better outcomes in patients when anticoagulant therapy is managed by an AMS or an AC compared with management by their personal physicians (ie, UC). Four such studies have reported major bleeding rates ranging from 2.8% to 8.1% per patient-year of UC.[203-206] Rates of thromboembolism with UC were not reported except in two studies in which the event rates were 6.2% and 8.1% per patient-year. Similarly, retrospective and prospective cohort studies of care provided by an AMS reported rates of major hemorrhage or thrombosis ranging from 1.4% to 3.3% and 0.7% to 6.3% per patient-year of therapy, respectively.[164,170,207-210] Three retrospective comparative studies using a before-and-after design of patients managed by UC or an AMS reported significant improvements in the outcomes of hemorrhage or thrombosis with AMS-directed care.[211-213] In contrast, however, two prospective, randomized controlled trials[214,215] comparing UC with the care of an AMS failed to show a significant difference in major hemorrhage or thromboembolism. The study by Matchar et al[214] also failed to show a significant improvement in TTR, although the AMS performed modestly better than UC. Wilson et al[215] did observe a significant improvement in TTR in the AMS group compared with UC (82% vs 76% respectively, $P = .034$). They also noted more high-risk INRs with UC than with an AMS (40% vs 30%, $P = .005$). This latter study had a major

limitation in that all patients were initially managed in an AMS for 3 months until they were stable and then observed for only 3 months after randomization to either receive UC or to continue care by the AMS.[215] The other study[214] suffered from a high turnover of patients, the possibility of selection bias affecting the referral of patients to the AMS, the open nature of the study, and targeted ranges that were sometimes outside recommended guidelines. In a retrospective, multicenter, international cohort study, 1,511 patients were randomly recruited from clinics offering routine medical care (UC) in the United States, Canada, and France, and from ACs in Italy and Spain.[216] Major management differences were detected, especially between AC care and routine medical care, but also among countries. For patients managed by routine medical care, documentation of care was often inadequate. Finally, less time in therapeutic INR range was noted in routine medical care.

In a systematic review of 67 studies representing > 50,000 patients managed by ACs (68%), during clinical trials (7%), or in community practices (24%), van Walraven et al[217] found that the practice setting had the greatest effect on anticoagulation control. TTR (days) varied from 56.7% in community practices to 65.6% in ACs and 66.4% for randomized trials. Compared with randomized trials the absolute reduction of TTR for community practices was −12.2% (95% CI, −19.5 to −4.8). The difference between community practices and ACs was −8.3% (95% CI, −4.4 to −12.1).

1.7.2. Computerized Monitoring: Computer assistance by the use of dedicated programs may improve dose management and TTR. Although programs differ, they typically calculate whether a dose adjustment is necessary from a user-defined table of trend rules for each therapeutic range. If dose adjustment is required, the current INR is compared with the target INR, and a proprietary equation calculates the new dose. The time to the next test is also set by the program using a set of variables comparing the current INR, the interval from the last test, the number of previous changes, and the number of previous INR values within the target range.

A number of early studies[218-220] evaluated computer programs to improve warfarin dosing. The first randomized study in 1993[221] showed that three contemporary computer programs all performed as well as the experienced medical staff of an anticoagulation management service in achieving a target INR of 2.0 to 3.0, but the computer achieved significantly better control when more intensive therapy was required (ie, INR, 3.0-4.5). In another randomized study[222] of 101 patients receiving long-term anticoagulation, computerized warfarin dose adjustment proved comparable to manual regulation in terms of the percentage of INR values maintained within the therapeutic range but required 50% fewer changes of dose. The first multicenter randomized trial, in 285 patients, of one computerized dosage program in 1998[223] showed a 22% overall improvement of control with the program compared with the performance of experienced medical staff. The computer program gave significantly better overall INR control across all target INR ranges. A slight improvement in TTR was also obtained by Italian investigators[224] using a different management program in > 1,200 randomized patients from five centers. A total of 71.2% of patients were in range with computer dosing and 68.2% were in range by manual dosing during the maintenance phase.[224] In both of these studies, the computer's improved results were probably due, in part, to a reduced propensity to reduce warfarin doses when the INR was at the upper limit of the "desired" range.

More recently, the European Concerted Action on Anticoagulation completed the first international, multicenter, randomized, controlled trial aimed to compare the safety and effectiveness of computer-assisted dosing of VKAs, using two computer programs in different centers (DAWN AC; 4S Information Sys, and PARMA; Instrumentation Laboratories), with conventional manual dosing.[225] A total of 13,052 patients treated with a VKA were followed for a total of 18,617 patient-years. Overall, there was a trend toward fewer clinical events with computer-assisted dosage. There was also a statistically significant reduction in clinical events in the subgroup treated for VTE. A subsequent cost-effectiveness analysis found computer-assisted dosing cheaper than manual dosing.[226]

The results of two small clinical trials suggest that computerized dose management can also be at least as effective as manual dosing by trained anticoagulation physicians or nurses during the initiation of anticoagulation therapy.[224,227] However, some computerized programs are unable to manage dosing during the induction phase.

1.7.3 Point of Care INR Testing: Point of care (POC) monitors measure a thromboplastin-mediated clotting time using a fingerstick sample of capillary whole blood or un-anticoagulated venous whole blood.[228] The result is then converted to a plasma PT equivalent by a microprocessor and is expressed as a PT and/or INR. Each manufacturer typically establishes the conversion formula by simultaneously comparing fingerstick or venous whole blood results with an established laboratory method and reagent that is traceable to the international reference thromboplastin.

Numerous studies[229-248] have reported on the accuracy and precision of these instruments and on the ability of patients, both adults and children, to obtain an INR and to use that INR to guide their anticoagulant therapy. However, limitations to the accuracy and precision of POC INR monitors have been documented. Problems identified with POC instruments include greater differences compared with a standard plasma-based methodology as INRs increase above the therapeutic range,[242,243] incorrect calibration of the ISI of the POC instruments,[244] the inability to calculate a mean normal PT,[249] and instrument-specific inaccuracies of the INR in patients with antiphospholipid antibodies.[250] In a recent systematic review of the literature, Gialamas et al[251] concluded that there is still no robust evidence that the use of POC instruments in general practice improves health outcomes or is cost-effective compared with UC and that analytical quality is comparable to laboratory testing. A major problem of comparative studies is the fact that there is a similar lack of correlation of INR results when anticoagulated plasmas are simultaneously compared using different instrument/thromboplastin combinations.[129-135] These differences may be clinically important in that they may lead to different dosing decisions.[128-134] Kaatz et al[252] compared two POC monitors and four clinical laboratories against a secondary reference thromboplastin preparation. They found that laboratories using a more sensitive thromboplastin showed close agreement with the standard, whereas laboratories using an insensitive thromboplastin showed poor agreement. The two POC monitors fell between these two extremes.

Steps are still needed to ensure the conformity of POC PT monitors to the WHO INR PT standardization scheme, but the WHO ISI calibration procedure is not practicable using the monitors. Simpler procedures for ISI calibration of POC monitors have recently been evaluated in a number of multicenter sites by the European Concerted Action on Anticoagulation and the UK National External Quality Assessment Schemes. By using lyophilized plasma calibrants with independently certified INRs, Poller and colleagues[253-255] have shown that verification or recalibration of the ISI of the instrument is possible. However, to obtain reliable ISI values for the two instruments tested they had to develop different ISI calibration methods. It is likely, therefore, that different types of POC monitor systems will require different ISI calibration methods. In a study of proficiency testing of three POC monitors over 6 years in >10 centers, Kitchen et al[256] found in each survey that INR results in 10% to 11% of centers were $>15\%$ different from results in other centers using the same monitors. This compared with a 12% difference for hospitals using conventional INR techniques. The European Concerted Action on Anticoagulation has recently tested 523 Coaguchek monitors at nine clinics and found that 20.3% of the monitors showed significant deviations.[257]

PST or PSM using a POC instrument represents another model of care with the potential for improved outcomes as well as greater convenience.[258] Several systematic reviews have shown improvements in the quality of anticoagulation control (TTR) and/or incidence of adverse events with PST and/or PSM.[259-261] Heneghan et al[261] pooled estimates from 14 randomized trials of PST showing a significant reduction in thromboembolic events (OR, 0.45; 95% CI, 0.30-0.68), all-cause mortality (OR, 0.61; 95% CI, 0.38-0.98), and major hemorrhage (OR, 0.65; 95% CI, 0.42-0.99) vs the comparator. For PST and PSM combined, there were significant reductions in thromboembolic events (OR, 0.27; 95% CI, 0.12-0.59) and death (OR, 0.37; 95% CI, 0.16-0.85) but not major hemorrhage (OR, 0.93; 95% CI, 0.42-2.05). In a recent randomized study from France comparing monthly laboratory monitoring with weekly self testing and monitoring in patients with prosthetic heart valves, self monitoring improved INR stability and reduced the incidence of bleeding, although the study was underpowered to show a true difference in clinical outcomes.[262] Another recent randomized study from Germany compared INR monitoring by the primary care physician with self management in patients with prosthetic heart valves and found greater INR stability and a lower incidence of thromboembolic events in the group of self-managed patients but no differences in bleeding events.[263]

PST and PSM require special patient training to implement.[264,265] This mode of therapy may not be suitable for all patients and may not be practical or cost-effective in certain settings. All participants in PST/PSM programs should participate in a recognized external quality assessment program.

1.8 Optimal Intensity Ranges

The optimal target range for the INR is not the same for all indications. In general, because bleeding is closely related to the intensity of anticoagulation,[164,266-268] there has been interest in establishing the lowest effective therapeutic range for each indication.[268-276]

Investigators have used various methodological approaches to establish the most appropriate range for different indications. These are as follows: (1) randomized trials in which patients are assigned to one of two different target ranges[270-274]; (2) indirect comparisons, in which outcomes are compared between separate randomized trials of VKA therapy that applied different target ranges of INR, and the control

patients received no therapy or another antithrombotic agent (usually aspirin)[277-280]; (3) subgroup analyses of observational studies (including within treatment groups of randomized trials) relating the observed INR or time spent in an INR range at the time of the outcome to either a bleeding event or thromboembolic event[164,207,208,267,268,281]; and (4) case-control studies in which the INR levels at the time of an event are recorded and compared with INR levels in appropriately selected control subjects.[174]

When moderate-intensity INR (approximately 2.0-3.0) was compared with higher-intensity adjusted-dose oral anticoagulation,[269,272-274,278-280] the moderate treatment intensity was shown to reduce the risk of clinically important bleeding without reducing efficacy. Conversely, a lower treatment intensity (eg, INR range 1.5-2.0) appears to be less effective than moderate-intensity therapy. For example, a randomized trial demonstrated that an INR of <2.0 (INR target, 1.5-2.0)[270] reduced the recurrence of venous thrombosis after an initial 3 to 6 months of standard treatment when compared with placebo. A subsequent clinical trial,[271] however, found that maintaining an INR intensity of 2.0 to 3.0 in the same clinical setting was more effective than a lower intensity of 1.5 to 2.0 and was not associated with a greater risk of bleeding. Likewise, in patients with AF, a randomized trial[282] reported that adjusted-dose warfarin therapy (INR, 2.0-3.0) was more effective than the combination of fixed-dose warfarin (3 mg/d) and aspirin; other studies showed that the efficacy of oral anticoagulant agents is reduced when the INR falls to <2.0.[174,207,208,283-285] Warfarin targeted to an INR of <2.0 has also been shown to be ineffective for preventing failure of dialysis access grafts.[286]

The use of fixed minidose warfarin (1 mg daily) has been evaluated in a number of clinical settings. A number of these studies reported that fixed minidose warfarin is ineffective when compared with dose-adjusted warfarin, in particular for the prevention of stroke in patients with AF and for the prevention of thrombosis of central venous catheters.[287-290] There is currently no evidence to support the use of fixed mini-doses of either acenocoumarol or phenprocoumon in any setting. However, the data obtained with fixed minidoses of warfarin cannot be extrapolated to other VKAs because of their different half-lives and the likelihood they would have different pharmacodynamic effects at similar doses. The optimal target range for each indication is discussed specifically in other articles in this supplement pertaining to each indication.

1.9 Significance of Nontherapeutic INRs

Fluctuations in INR may occur because of any one or more of the following conditions: inaccuracy in INR testing, changes in vitamin K intake, changes in the absorption of vitamin K or VKAs, changes in the metabolism of VKAs, changes in vitamin K-dependent coagulation factor synthesis or metabolism, other effects of concomitant drug use, or patient noncompliance. A number of studies have shown that adverse event rates rise sharply as the INR moves above or below the target INR interval.[163,164,174] A recent retrospective analysis of $>3,000$ patients with AF found that the one-third with the poorest INR control (48% of time in range) had twice the rate of stroke, myocardial infarction, major bleeding, and death as did the one-third with the best INR control (83% of time in range).[163]

However, the risk of adverse events associated with a single INR outside the therapeutic range is probably low. For example, although excessively elevated INR values are clearly associated with an increased risk of bleeding, in particular for INR values of >5.0,[164,169,170,268] data from a large registry of warfarin-treated patients suggest that the short-term risk for major bleeding is low for someone with a single INR value between 5.0 and 9.0 (0.96% at 1 month).[291] Similarly, in a retrospective, matched cohort study of 2,597 patients on warfarin therapy, the risk of thromboembolic events at 3 months in patients with stable INRs who experienced a single significant subtherapeutic INR value was low (0.4%) and was not significantly different from the risk observed in patients with persistently stable INRs.[292] Similar rates were reported in a population of 294 patients with mechanical heart valves experiencing a single subtherapeutic INR and in a large cohort of patients who required short-term discontinuation of warfarin to undergo minor outpatient interventions.[293,294]

Before executing a plan for managing an episode of altered anticoagulation effect, important factors relating to the collection and processing of the blood sample must be taken into consideration. For example, a spuriously elevated INR should be suspected when a patient with previously stable INR control presents with a very high INR result in the absence of any explanation for the loss of INR control. Spurious results might also be suspected when abnormalities were encountered during sample collection (eg, the phlebotomist had a difficult time obtaining a sample or combined the contents of two collection tubes, and so forth).[295] Recommendations for the management of patients whose INR is outside the therapeutic range are provided in Holbrook et al.[86]

1.10 Adverse Events

1.10.1 Bleeding Events: The rate of hemorrhagic events must be interpreted in the context of the clinical characteristics of the group studied. Factors that

influence the rate of bleeding include the following: the target INR range; whether patients are mostly new to therapy or have long-term experience with therapy; whether an INR or PT is used to manage therapy; the indication for anticoagulation; the type of VKA; patient-specific risk factors, including concomitant antiplatelet therapy; and the quality of dose management. It is also not appropriate to extrapolate the rates of adverse events from randomized controlled trials to everyday practice, because high-risk patients may be excluded from clinical trials, and monitoring and management of anticoagulation are often better coordinated in clinical trials than in clinical practice.

When bleeding occurs, especially from the GI or urinary tract, the presence of an underlying occult lesion should always be considered. This is important because the patient factor that most consistently predicts major bleeding is a history of other bleeding, particularly from the GI tract.[207] A number of descriptive studies[296-298] have reported on the probability of finding occult lesions. Coon and Willis[296] identified occult lesions that were responsible for bleeding in 11% of 292 patients with hemorrhage. Jaffin et al[297] found a 12% prevalence of positive stool occult blood test results in 175 patients receiving warfarin or heparin compared with 3% in 74 control subjects. There was no difference between the mean PT or activated partial thromboplastin time (aPTT) in patients with positive and negative test results. Among the patients with positive stool occult blood test results, 15 of 16 patients had a lesion that had not been previously suspected, and four patients had neoplastic disease. Landefeld et al[266] found that 14 of 41 patients with GI bleeding had important remediable lesions, of which two were malignant. To perform endoscopy in these patients is an important choice, because endoscopy has been shown in prospective studies to identify the bleeding source in $>50\%$ of patients managed with anticoagulant therapy who present with upper GI tract bleeding and because endoscopic therapy for nonvariceal upper GI bleeding achieves hemostasis in $>90\%$ of patients.[299] This limited information supports the need to investigate patients with occult GI bleeding, as it may herald the presence of an underlying malignancy or other lesion that is frequently rectifiable.

In a 2-year prospective study in which enrolled patients had monthly urinalysis, Culclasure et al[300] found microscopic hematuria in 3.2% of patients receiving oral anticoagulation compared with 4.8% in the control group not receiving anticoagulant therapy. There was no difference in the rate of hematuria with therapeutic or high INRs. Following a second episode of hematuria, 43 patients (32 receiving anticoagulant therapy, 11 control patients) were investigated. Of these patients, 27 receiving anticoagulant therapy (84%) and eight control patients (73%) were found to have significant underlying disease, with three cancers found in the combined group (7%). Other case series[301,302] have reported a higher likelihood of underlying lesions in patients who develop hematuria while receiving anticoagulant therapy.

1.10.2 Factors Predictive of Bleeding Events: The most important factor influencing the risk of bleeding is the intensity of anticoagulant therapy.[162,164,169,170,174,208,266-269,272-275,283-285] The likelihood of bleeding has been reported to rise steeply as the INR increases above 5.0.[164,169,170,268] The optimal target range for each indication and the lowest effective range are discussed specifically in other articles in this supplement pertaining to each indication.

Several patient characteristics are associated with higher odds of bleeding during anticoagulation therapy. The patient factor most consistently predictive of major bleeding is a history of bleeding (especially GI bleeding).[162,208,273] Other factors associated with a higher risk of bleeding include advanced age; the presence of a serious comorbid condition, such as cancer, renal insufficiency, liver disease, arterial hypertension, and prior stroke; alcohol abuse; and the use of concomitant therapies, in particular antiplatelet drugs.[162,203,204,207,303]

A number of prediction models of bleeding risk have been proposed. Dahri and Loewen[304] performed a qualitative review of the published clinical prediction rules that estimated bleeding risk in patients starting on warfarin. The authors found seven studies, of which four presented distinct clinical prediction rules. Because none of these scores exhibited sufficient predictive accuracy or evaluated the impact of their use on patient outcomes, the authors concluded that no existing clinical prediction rules can be recommended for widespread use in practice at present. One prediction model that was prospectively validated in different outpatient populations identified four independent factors associated with an increased risk of bleeding. These factors were age >65 years, history of gastrointestinal bleeding, history of stroke, and at least one of the following variables: myocardial infarction, hematocrit $<30\%$, creatinine >1.5 mg/dL, and diabetes.[203] Another model, derived using registry data from patients with AF, identified prior bleeding, hepatic or renal disease, ethanol abuse, malignancy, age >75 years, reduced platelet count or function, hypertension (uncontrolled), anemia, genetic factors (CYP2C9 polymorphism), an excessive risk of falls, and a history of stroke as risk factors for bleeding.[305]

The impact of age on bleeding risk remains controversial, with older reports finding risk increasing with age,[162,163,181,267,306,307] whereas newer studies have failed

to find this association.[170,205,207,268,282,308-310] The discrepancy may be partly explained by the wide range in the mean age of the patients enrolled in the various studies, the relative lack of representation in most studies of patients > 80 years of age, and the selection and survivorship biases in noninception cohort studies. When investigators attempt to separate the effect of age from comorbid conditions associated with age, some have concluded that age in and of itself is not a major independent risk factor,[112,267,311] whereas others have found it to be an independent risk factor[266,268] even after controlling for the intensity of the anticoagulant effect. Some have suggested that older patients may have a lower risk of bleeding when managed by AMS.[312,313] Even if the overall risk of bleeding is not increased in the elderly, it is clear that the risk of intracranial hemorrhage increases with age.[267,268,314,315]

Warfarin is frequently used concurrently with other antithrombotic agents. A meta-analysis assessing clinical studies comparing warfarin alone and warfarin in association with aspirin found that the combination increased the risk of bleeding by almost one-half compared with warfarin alone (OR, 1.43; 95% CI, 1.00-2.02).[316] A combined analysis of the SPORTIF (Stroke Prevention Using an Oral Thrombin Inhibitor in Atrial Fibrillation) III and V trials involving 481 patients who received aspirin and warfarin compared with 3,172 patients who received warfarin alone demonstrated a significant 1.6% per year increase in major bleeding.[317] Although there are no randomized controlled trials that have compared bleeding rates in patients receiving "triple therapy" (usually warfarin, aspirin, and clopidogrel) with either warfarin alone or with a "dual therapy," a systematic review identified 12 reports involving 3,413 patients treated with oral anticoagulants who underwent percutaneous coronary intervention with stent insertion and subsequently received the combination of aspirin, clopidogrel, and warfarin. The rates of major bleeding in patients receiving triple therapy ranged from 0% to 21% (mean 7.4%) during up to 21 months of follow-up and 0% to 5.9% (mean 2.6%) during 30 days of follow-up.[318] In a Danish nationwide registry of patients with AF, all combinations of warfarin, aspirin, and clopidogrel were associated with an increased risk of nonfatal and fatal bleeding, whereas dual or triple therapy carried a more than threefold higher bleeding risk than warfarin alone.[319]

1.10.3 Reversal Strategies: Strategies to reverse the effect of VKAs may be needed in patients who require urgent invasive procedures, in asymptomatic patients presenting with excessively elevated INR values, and in bleeding patients. Therapeutic options include interruption of VKA treatment as well as the administration of vitamin K (usually vitamin K1, phytonadione) and blood derivatives such as fresh frozen plasma and prothrombin complex concentrates and recombinant activated factor VII.[303,320]

Interruption of VKAs may be sufficient in patients who need an elective invasive procedure or in asymptomatic patients with an elevated INR value and a low risk for bleeding. In this latter case, it must be noted that it takes approximately 2.5 days for an INR between 6.0 and 10.0 to decline to < 4.0.[321] Because acenocoumarol has a much shorter half-life than warfarin, the time required for an effective decline will be less and is probably no more than 1 day for most patients.[322,323] Conversely, the longer half-life of phenprocoumon will probably result in a much slower decline. Finally, the half-life of fluindione is similar to that of warfarin, and thus a similar decline should be expected.

Phytonadione (vitamin K_1, a form of vitamin K derived from plants) has been used in clinical trials that assessed the usefulness of oral vitamin K for the treatment of warfarin-associated coagulopathy.[324] Vitamin K_2 (menaquinone, which is synthesized by bacteria) and vitamin K_3 (menadione) are not widely available and have not been well studied in clinical trials.[324] Low doses of phytonadione can be administered orally in combination with warfarin interruption in patients not requiring urgent reversal. When oral phytonadione is administered in conjunction with temporary interruption of warfarin therapy, approximately 1.4 days are required for an INR between 6 and 10 to decline to < 4.0.[321] When administered intravenously, low doses of phytonadione produce similar reductions as oral phytonadione in the INR value at 24 h, whereas subcutaneous phytonadione appears to be less effective than low-dose oral phytonadione.[325,326] When administered at higher doses for the management of the bleeding patient, intravenously administered phytonadione works more rapidly than either oral or subcutaneous vitamin K_1.[325,327,328] Reduction of the INR begins within 2 h, and a correction to within the normal range is generally achieved within 24 h if hepatic function is normal and if a sufficiently large dose is given.[324] At 24 h, 5 mg of oral and 1 mg or IV vitamin K_1 produce similar effects on the INR.[325] IV phytonadione may cause anaphylactoid reactions. Although frequently reported, and likely more common in patients who receive large IV doses administered rapidly, the true frequency of this complication is about three per 10,000 doses administered, and it may be more likely to occur if formulations containing polyethoxylated castor oil are used to maintain the vitamin K in solution.[329] To minimize the risk of anaphylactoid reactions, vitamin K_1 should be mixed in a minimum of 50 mL of intravenous fluid and administered, using an infusion pump, over a minimum of 20 min.

Fresh frozen plasma remains the most widely used coagulation factor replacement product for urgent reversal of coumarin anticoagulation.[320] Plasma may be a potential carrier of infective agents, and its use is associated with an increased risk of volume overload. Furthermore, it requires a cross-match if group-specific plasma is to be used, and it takes a prolonged period of time to thaw and administer. Given the long half-life anticoagulant effect of warfarin and the short half-life of infused coagulation factor concentrates, phytonadione must also be given to restore the adequate endogenous production of VKA-sensitive anticoagulant proteins. Urticaria occurs frequently with plasma transfusion; anaphylaxis is less common, occurring in about one in 20,000 transfusion episodes.[330] Transfusion-related acute lung injury remains the most feared complication after transfusion and is estimated to occur in about one in 5,000 plasma-containing transfusions.[331]

Nonactivated prothrombin complex concentrates (PCC) are probably more effective than plasma in correcting INR. PCCs do not require a cross-match, are virally inactivated, do not pose a risk of volume overload, and can be infused in 15 to 30 min. PCC may be classified as three-factor products (with adequate levels of factors II, IX, X, and low factor VII levels) and four-factor products, which contain adequate levels of factors II, VII, IX, and X as well as protein C and S.[319,331] Current PCCs are more or less devoid of activated clotting factors and are supplemented with heparin and antithrombin to minimize the risk of thrombosis.[332] Four-factor PCCs are currently not available in some countries (eg, United States).

In patients with life-threatening bleeding, recombinant activated factor VII has been used to control bleeding.[320] Recombinant activated factor VII is able to generate a consistent thrombin burst through both tissue factor-dependent and tissue factor-independent mechanisms and is able to trigger thrombin generation even in the presence of significant platelet dysfunction. Evidence supporting its use in VKA-associated bleeding is currently limited, and its use cannot be recommended except in the setting of life-threatening bleeding when more effective agents are not available.[333] As would be expected based on its potent procoagulant effect, this agent may cause thrombosis. Recommendations to guide the use of these strategies are provided in Holbrook et al.[86]

1.10.4 Nonhemorrhagic Adverse Events: Other than hemorrhage, the most important side effects of warfarin are acute thrombotic complications, such as skin necrosis and limb gangrene. These uncommon complications are usually observed on the third to eighth day of therapy[334,335] and are caused by extensive thrombosis of the venules and capillaries within the subcutaneous fat (in the case of skin necrosis) and massive outflow obstruction of the venous circulation of the limb (in the case of limb gangrene). The pathogenesis of these complications and the reason for the localization of the lesions are not well understood. An association between warfarin-induced skin necrosis and protein C deficiency[336-339] and, less commonly, protein S deficiency[339] has been reported, but this complication also occurs in nondeficient individuals. A pathogenic role for protein C deficiency is supported by the similarity of the lesions to those seen in neonatal purpura fulminans, which complicates homozygous protein C deficiency. A variant of this syndrome also attributed to a severe warfarin-induced depletion of protein C is the occurrence of venous limb gangrene during warfarin treatment of cancer-associated DVT[340] and in some patients with heparin-induced thrombocytopenia started on warfarin after withdrawal of heparin.[341,342] The management of patients with warfarin-induced skin necrosis who require lifelong anticoagulant therapy is problematic. Therapy with warfarin is considered to be contraindicated, and long-term heparin therapy is inconvenient and is associated with osteoporosis. A reasonable approach in such patients is to restart warfarin therapy at a low dose (eg, 2 mg), under the coverage of therapeutic doses of parenteral anticoagulants, and to gradually increase the warfarin dose over 1 or more weeks. This approach should avoid an abrupt fall in protein C levels before there is a reduction in the levels of factors II, IX, and X, and it has been reported to not be associated with the recurrence of skin necrosis in a number of case reports.[337,338,342]

The purple toe syndrome may very rarely occur in association with the initiation of VKA treatment. It is a nonhemorrhagic, cutaneous complication due to cholesterol emboli that usually develops 3 to 8 weeks after the start of warfarin therapy and is characterized by the sudden appearance of bilateral, painful, purple lesions on the toes and sides of the feet that blanch with pressure.[343,344]

The VKAs also interfere with the carboxylation of Gla proteins that are synthesized in bone.[345-348] Although these effects contribute to fetal bone abnormalities when mothers are treated with a VKA during pregnancy,[349,350] it is unclear how they might affect children. There are two uncontrolled cohort studies that describe reduced bone density in children on warfarin for >1 year, but the role of the underlying disorders in reducing bone density remains unclear.[351]

Finally, by preventing the activation of G1a proteins and growth arrest-specific gene 6 (Gas-6), VKAs may also induce vascular calcification.[352] However, this relationship in humans is conjectural and is based on case reports.[352]

2.0 Direct Thrombin Inhibitors: Dabigatran Etexilate

Dabigatran is a selective, reversible, direct thrombin inhibitor given as dabigatran etexilate, an orally absorbable prodrug, since dabigatran itself is a strongly polar molecule that is not absorbed from the gut. Phase 3 clinical studies reported to date have evaluated the use of dabigatran etexilate for the prevention of VTE after elective total knee or hip arthroplasty, for therapy of VTE, and to prevent stroke or systemic embolism in nonvalvular AF. The drug is approved in many countries for the prevention of VTE in patients undergoing total hip or knee replacement surgery and in the United States and Canada for the prevention of stroke or systemic embolism in nonvalvular AF. Dosing schedules are 150 mg and 220 mg once daily when used to prevent VTE (starting with a half dose given soon after surgery) and 110 mg and 150 mg bid in patients with AF, although only the latter dose was approved for this use in the United States. The 150 mg bid dose was also used for the treatment of VTE. Dabigatran etexilate remains under evaluation for the secondary prevention of VTE and in the management of acute coronary syndromes. Melagatran (given as its prodrug Ximelagatran) was the first of the orally available direct thrombin inhibitors to be clinically evaluated and was effective for both the prevention and treatment of VTE and in AF but caused unacceptable liver toxicity.[353-355]

2.1 Pharmacology

Direct thrombin inhibitors, such as dabigatran, the hirudins, and argatroban, do not require a cofactor, which differentiates them from the indirect coagulation inhibitors like the heparins, other glycosaminoglycans, and the synthetic pentasaccharide that must form a complex with plasma antithrombin before they can accelerate inhibition of thrombin and/or factor Xa.[356] Clot-bound thrombin is relatively protected from heparin-like anticoagulants in vitro but is freely accessible to direct thrombin inhibitors.[119]

Dabigatran (initially referred to as BIBR 953) is a small synthetic molecule of 471.5 d with strongly basic functional groups that is a concentration dependent, competitive, highly selective and reversible direct thrombin inhibitor with a Ki of 4.5 ± 0.2 μM (which is comparable with that of melagatran).[357,358] The inhibitor prevents access to the active site of thrombin by forming a salt bridge between its amidine group and Asp 189 and through hydrophobic interactions.[357] Like melagatran and argatroban, dabigatran is a univalent inhibitor that interacts with the active site of thrombin alone, whereas hirudin, lepirudin, and desirudin also bind to a substrate recognition site (exosite 1).[356] Because the highly polar and lipophobic dabigatran molecule is not absorbed from the gut, its oral availability required the synthesis of an absorbable prodrug, dabigatran etexilate (BIBR 1048; 627.7 d).[357] On absorption, this esterified, hexylated, and more lipophilic prodrug is rapidly converted to dabigatran (the mechanism is ester cleavage catalyzed by serine esterase enzymes, via two intermediates—BIBR 1087E and BIBR 951 CL).[357,359] As would be expected from a direct thrombin inhibitor, dabigatran prolongs the thrombin clotting time (TCT), PT, aPTT, and ecarin clotting time (ECT) of plasma from humans, rats, rabbits, dogs, and rhesus monkeys and also inhibits thrombin generation in human plasma.[358] The ECT has been a preferred measure of anticoagulant effect for r-hirudin and other direct thrombin inhibitors; ecarin is a metalloprotease enzyme obtained from venom of the saw-scaled viper (*Echis carinatus*) that generates meizothrombin from prothrombin.[360,361]

Dabigatran prevents thrombin-induced platelet aggregation but not platelet aggregation by arachidonic acid, collagen, or adenosine diphosphate.[358] Administration of oral dabigatran etexilate or IV dabigatran causes concentration-dependent reductions of thrombosis provoked in rats and rabbits by venous stasis plus tissue factor infusion or venous stasis plus endothelial damage. Thrombus inhibition peaks within 30 to 60 min of an oral dose, then persists for 2 to 3 h in rats and about 7 h in rabbits, and correlates well with prolongation of the aPTT.[362,363] In rats, the intravenous dabigatran dose required to prolong the tail bleeding time is 5 to 10 times greater than the maximally effective antithrombotic dose of 0.1 mg/kg.[363]

2.2 Pharmacokinetics and Pharmacodynamics

Dabigatran etexilate is now formulated as a capsule containing multiple small pellets, each of which is composed of drug coated on a tartaric acid core to create an acid microenvironment that favors drug dissolution and preserves gut absorption even when the gastric pH is high (solubility is best at a low pH).[364] Systemic bioavailability of oral dabigatran etexilate has been measured at 7.2% in healthy young volunteers and estimated at 6% to 7% in healthy older subjects.[359,364] Oral bioavailability of the capsules is comparable to that of tablets used in the phase 1 and early phase 2 evaluations.[365,366] Plasma dabigatran concentrations peak within 2 hours after a dose of dabigatran etexilate and then decrease by $>70\%$ during an initial 4- to 6-h distribution phase that is followed by a much slower elimination phase. With repeated dosing, the terminal elimination half-life is 12 to 17 h, the peak and trough concentrations are dose-proportional, and it takes 2 to 3 days to reach

steady-state levels. The terminal half-life following a single dose is about 9 h in healthy volunteers.[359,361,364] The summary of pharmacokinetic parameters in Table 3 was derived from Stangier.[367] Rapid conversion of dabigatran etexilate to dabigatran ensures that plasma concentrations of the etexilate and two intermediate prodrugs barely reach detectable levels.[359] Approximately 35% of circulating dabigatran is protein-bound, regardless of concentration. After a 5-mg dose of IV dabigatran in healthy volunteers, the distribution volume was measured at 69 to 90 L, which exceeds the volume of body water, and plasma clearance was 149 mL/min.[359] About 15% of available dabigatran is conjugated to form pharmacologically active but unstable glucuronides that account for about 20% of the total drug exposure.[359] Eighty-five percent of the dose is excreted by renal clearance, almost all as unchanged dabigatran.[359] Pharmacokinetic data from the phase 2 studies, together with population modeling, predict average steady-state peak and trough plasma dabigatran concentrations of 99 and 14 ng/mL after daily dosing with 150 mg dabigatran etexilate,[368] 183 and 37 ng/mL after 220 mg/d,[369] and 184 and 90 ng/mL after 150 mg bid (Table S3).[369] Apart from dose, the systemic exposure to dabigatran is related to age and renal function. When older but apparently healthy volunteers aged 65 to 87 years received 150 mg dabigatran etexilate bid, the steady-state area under curve (AUC) was 1.7 to 2 times greater than that observed in a previous study of younger men aged 18 to 45 years. The increase in drug exposure was explained by reduced renal clearance and was more obvious in older women (in whom the AUC was 3% to 19% above that in older men).[364,367] In a separate parallel group volunteer study, the AUC∞ after 150 mg dabigatran etexilate was 1.5, 3.2, and 6.3 times higher in people with mild, moderate, or severe renal impairment (creatinine clearance of 51-80, 31-50, and ≤ 30 mL/min) than in healthy control subjects. The corresponding levels of peak plasma concentration (Cmax) were 109, 138, and 205 ng/mL, compared with 85 ng/mL when renal function was normal. The terminal half-life was doubled to 28 h in severe renal failure, from 14 h in control subjects.[370] Strong effects of renal function on drug concentrations were also demonstrated in patients having a hip or knee replacement, in whom dose-effect modeling predicted a steady-state Cmax of 100 ng/mL during bid dosing with 150 mg if the creatinine clearance was >90 mL/min, increasing to 140, 180, and 240 ng/mL as creatinine clearance diminished to 60 to 90, 40 to 60, and <40 mL/min.[366,367] Moderately severe liver dysfunction (Child-Pugh classification B) appears to have little effect on dabigatran pharmacokinetics, since peak plasma concentrations were reduced by 15% after 150 mg dabigatran etexilate in twelve affected subjects, when compared with 12 healthy age- and sex-matched control subjects, whereas time to Cmax, the elimination half-life, AUC, distribution volume, and extent of glucuronidation remained unchanged. Effects on blood clotting test results were similar in the two study groups.[371] The effects on blood coagulation tests closely mirror plasma dabigatran concentrations. Peak prolongation coincides with the Cmax, and clotting times decrease as dabigatran leaves the circulation. Unlike warfarin and other vitamin K inhibitors, which have long-lasting effects on the INR, the relatively short half-life of dabigatran means that time between dosing and blood sampling is a critical determinant of drug effect. Effects on the aPTT, INR, TCT, and ECT were studied in healthy volunteers and in patients having a hip or knee replacement. In the placebo-controlled volunteer study, wherein healthy men aged 18 to 45 years received one dose of 10 to 400 mg dabigatran etexilate or eight hourly doses of 50 to 400 mg for 6 days, the

Table 3—*[Section 2.2] Pharmacokinetic Characteristics of Dabigatran Etexilate*

Population and Dosage	Tmax or Tmax, ss, h	Cmax or Cmax, ss, ng/mL/mg	Cmin or Cmin, ss, ng/mL/mg	AUC∞ or ss, (ng/h/mL)/mg	$t_{1/2}$, h
Healthy adults					
od	1.25-1.5	0.89		5.66	8.13
bid	1.5-2.0	1.16	0.36	7.4	11.3
tid		1.68	0.50	7.85	13.7
Healthy elderly					
bid (men)	3.0	1.48	0.52	10.9	12.1
bid (women)	2.5	1.83	0.52	12.8	13.4
Patients					
od	6.0[a]	0.41[a]	0.13	6.41[a]	
bid	2.7	1.06	0.45	15.9	

The results recorded for bid or tid dosing are at ss. Cmax, Cmin and AUC are normalized to a 1-mg dose of dabigatran etexilate. Results are mean or median. AUC = area under the curve; AUC∞ = AUC from time zero to infinity; Cmax = peak plasma concentration; Cmin = minimum plasma concentration; od = once daily; ss = steady state; tmax = time to reach the peak plasma concentration; $t_{1/2}$ = terminal half-life. (Modified with permission from Stangier et al.[364])

[a]Measured after first oral dose of dabigatran etexilate following surgery for hip or knee replacement.

drug was given as a dry powder dissolved in dilute tartaric acid[361] and caused dose-dependent prolongations of aPTT, INR, TCT, and ECT. The coagulation test results correlated with plasma dabigatran concentration, the maximum clotting times coincided with Cmax, and effects persisted for 8 h or longer except for the aPTT and INR at the lowest doses. The peak clotting time ratios after 200 to 400 mg of dabigatran etexilate were >55 for the TCT, >5 for the ECT, 1.5 to 1.9 for the INR, and 1.8 to 2.1 for the aPTT. The TCT, ECT, and INR increased in linear proportion with plasma dabigatran levels, but the aPTT concentration-response curve was curvilinear and flattened above concentrations of 200 ng/mL.[361,369] Plasma dabigatran levels also correlated closely with effects on the ECT and aPTT in the BISTRO trial, a safety study of dabigatran etexilate in patients having an elective hip or knee arthroplasty, wherein escalating doses of dabigatran etexilate were given for 6 days in tablet form after surgery. The doses were 12.5 mg rising to 300 mg bid or 150 mg to 300 mg once daily.[368] Bid dabigatran etexilate raised the peak aPTT and ECT ratios from 1.08 to 1.91 and from 0.96 to 5.17 as the dose increased from 12.5 mg to 300 mg, whereas the mean trough levels of aPTT ratio rose from 1.0 to 1.65.[368] Concentration-effect relationships were linear for the ECT but nonlinear for the aPTT, with better sensitivity and precision for the ECT.[366] A nonlinear dose-effect on the aPTT was also found in patients with AF taking bid doses of up to 300 mg.[372] Different aPTT reagents appear to have similar sensitivities to dabigatran, as does the activated clotting time.[369] The likely impact of reduced renal function on blood clotting test results in patients given 150 mg dabigatran etexilate bid was estimated through population modeling of data from the BISTRO I study. This predicted a maximum aPTT of 47.4 s with normal renal function, increasing to 54.2, 61.9, and 78.3 s in patients with mild, moderate, or severe reductions in creatinine clearance (the corresponding predictions for maximum ECT were 55.2, 77.4, 108, and 183 s).[367] The modeling is consistent with results when otherwise healthy volunteers with a creatinine clearance <30 mL/min received a 150-mg dose of dabigatran etexilate: maximum aPTT, ECT, INR, and TCT ratios were 3.45, 3.26, 4.14, and 12.5, compared with 1.85, 1.89, 1.4, and 6.87 in healthy control subjects.[370] This emphasizes the need to avoid dabigatran etexilate in patients with severe renal failure and the importance of renal function when considering the choice of dosing regimen. It remains possible that dabigatran may activate platelets in patients with AF, as urinary excretion of 11-dehydrothromboxane B2 (an end product of thromboxane A2 metabolism) was increased by roughly 20% during 12 weeks of treatment in a dose-ranging trial, and that effect was suppressed by aspirin.[372]

2.3 Interactions

2.3.1 Factors Interfering With Absorption: Absorption of dabigatran etexilate is influenced by gastric pH as affected by proton pump inhibitors, food, the postoperative state, and also by drugs that inhibit or induce activity of the cell efflux transporter P-glycoprotein (P-gp). Dabigatran etexilate has a low aqueous solubility that is further reduced by increased pH,[373] as is observed in patients taking the gastric proton pump inhibitor pantoprazole. Twice-daily pretreatment with 40 mg pantoprazole for 48 h in a crossover study reduced geometric mean levels of Cmax and AUC after 150 mg dabigatran etexilate by 40% and by 32% in healthy male volunteers aged 18 to 55 years.[365] Bioavailability (steady-state AUC) was also reduced by 20% to 40% in a parallel group study in which older volunteers, aged ≥65 years, took 40 mg pantoprazole with 150 mg dabigatran etexilate bid for 6 days; pantoprazole raised the gastric pH from 2.2 to 5.9, and pH correlated with AUC. Small corresponding changes in Cmax, ECT, and aPTT were not believed to have clinical importance.[364]

Taking 150 mg dabigatran etexilate after a high-fat, high-calorie breakfast prolonged the time taken to reach Cmax from 2 h to 4 h in the crossover study described above, although Cmax and total drug exposure remained unchanged.[365] Absorption of doses taken 4 to 8 h after a hip replacement was slowed and reduced, compared with 2 to 10 days later, such that time to reach the peak plasma concentration was delayed to 6 h and both Cmax and AUC were greatly diminished (Table 3); the changes were attributed to early effects of surgery on GI motility and gastric acidity.[366]

2.3.2 Other Drug-Drug Interactions: Important drug-drug interactions most often result from changes in drug metabolism that are due to induction or inhibition of CYP3A4 and other enzymes of the microsomal cytochrome P450 complex or from changes in drug bioavailability mediated by the adenosine diphosphate-dependent cell efflux transporter, P-glycoprotein (P-gp).[374] Potential drug interactions with dabigatran etexilate have been explored in studies in which volunteers received dabigatran etexilate together with drugs known to provoke such mechanisms. Many drugs may interact through more than one pathway.

Because cytochrome P450 enzymes have almost no role in the metabolism of dabigatran and are not affected by dabigatran in vitro, this becomes an unlikely mechanism for drug-drug interactions,[359] and volunteer studies confirm the lack of a clinically important interaction with atorvastatin (a substrate for CYP3A4 and substrate/inhibitor of P-gp) and diclofenac (a substrate for CYP2C9 and uridine

glucuronyltransferase 2b7, and also a substrate and weak inhibitor of UGT1A).[367] When 22 volunteers aged 43 ± 15 years took 80 mg atorvastatin together with 150 mg dabigatran bid for 4 days in an open-label crossover study, the steady-state AUC of dabigatran was reduced by 18%, whereas the Cmax and AUC of atorvastatin increased by 15% and 23%.[375] Cmax and AUC of dabigatran remained unchanged in a similar study of 24 volunteers aged 18 to 55 years who took one 50-mg dose of diclofenac after 4 days of bid dosing with 150 mg dabigatran etexilate, whereas Cmax of diclofenac and its main metabolite decreased by 11% to 17%.[376] The changes were believed to be small and clinically unimportant.

In vitro studies find that dabigatran etexilate (but not dabigatran) is a substrate for P-glycoprotein (P-gp, MDR1) with a medium affinity when tested using the Caco-2 cell-line1, which makes it a potential target for P-gp-related drug interactions.[377] The bioavailability of P-gp substrates like dabigatran may be raised or reduced through inhibition or induction of P-gp: the P-gp inhibitors include amiodarone, verapamil, ketoconazole, quinidine, and clarithromycin, whereas P-gp inducers include rifampicin and St. John's wort (*Hypericum perforatum*).[377]

In formal interaction studies with amiodarone, a first dose of 600 mg raised the AUC and Cmax of dabigatran by about 50% and 60%, an interaction that may persist for some weeks after stopping amiodarone due to the long half-life of this drug.[373] The effects of verapamil depend on its dosing schedule and drug formulation.[373] The first dose of an immediate-release formulation, when given 1 hour before 150 mg dabigatran etexilate, increased the Cmax and AUC of dabigatran by about 180% and 150%, but these elevations were reduced to about 60% and 50% after repeated dosing and to about 90% and 70% when taking an extended-release formulation. The interaction became negligible (increases of 10% in Cmax and 20% for AUC) if verapamil was taken >2 h after dabigatran etexilate when dabigatran absorption was essentially complete. Twice-daily coadministration of 500 mg clarithromycin increased the AUC and Cmax of dabigatran by about 19% and 15%, respectively.

The strong P-glycoprotein inhibitors quinidine and ketoconazole are contraindicated when taking dabigatran etexilate because they markedly increase exposure to dabigatran.[373] The approved product information advises caution when considering coadministration of strong P-gp inducers like rifampicin or St. John's wort, which may significantly decrease the Cmax and AUC.[373]

The P-gp substrate digoxin is used to probe P-gp mediated drug-drug interactions. After 4 days of dosing with once-daily digoxin plus bid 150 mg dabigatran etexilate, there was little effect on the pharmacokinetics of either drug in a three-way crossover study of 23 healthy volunteers aged 18 to 65 years.[378]

Drug interactions that may change the Cmax and AUC of dabigatran have not been correlated with clinical outcomes. Very large increases of Cmax and AUC, like those of quinidine and systemic ketoconazole, are likely to raise the bleeding risk. Moderate increases, like those provoked by amiodarone or verapamil, may become important if combined with old age or reduced renal clearance. The concern about strong P-gp inducers like rifampicin is their potential to decrease drug exposure and therefore reduce efficacy.

2.3.3 Dabigatran Etexilate and Antiplatelet Drugs: The added bleeding risk when platelet function inhibitors like aspirin and clopidogrel are taken during anticoagulant therapy is compounded for aspirin and other nonsteroidal antiinflammatory drugs by the increased likelihood of peptic ulceration due to interference with prostaglandin-mediated cytoprotection of the gastrointestinal mucosa.[379] These mechanisms are independent from any pharmacokinetic drug-drug interactions (none was demonstrated between dabigatran etexilate and diclofenac).[376] Aspirin increased the bleeding rate when added to ximelagatran in patients with AF and to warfarin in patients with AF, a prosthetic heart valve, coronary artery disease, or peripheral vascular disease.[82,317,380]

The added effects on bleeding when aspirin is combined with dabigatran etexilate were explored in the Prevention of Embolic and Thrombotic Events in Patients With Persistent Atrial Fibrillation (PETRO) trial, a phase 2, parallel-group, randomized, dose-ranging safety study in 502 patients with nonvalvular AF who also had coronary artery disease and/or one or more other risk factors for systemic embolism. The patients received 12 weeks of treatment with open-label warfarin alone (target INR 2-3) or with blinded 50 mg, 150 mg, or 300 mg dabigatran etexilate bid plus daily aspirin (81 mg or 325 mg) or a placebo, using a 3 × 3 factorial design to allocate study treatments to patient groups of unequal size. Aspirin increased the chances of major or clinically significant nonmajor bleeding in patients given a supratherapeutic dabigatran dose of 300 mg bid, in whom the bleeding rate was 20% with an aspirin dose of 325 mg/d (six of 30), 14.7% if the dose was 81 mg/d (five of 34), and 5.7% (six of 105) in patients given the aspirin placebo. The trends reached statistical significance when the two aspirin groups were pooled. Bleeding risk was not apparently raised when aspirin was added to 50-mg or 150-mg doses of dabigatran, but the sample size was too small to exclude clinically important effects.[372]

Low doses of aspirin (≤ 100 mg/d) were permitted in the Randomized Evaluation of Long-term Anticoagulant Therapy (RE-LY) (nonvalvular AF) and Dabigatran in the Treatment of Venous Thromboembolism (RE-MEDY) (VTE) phase 3 studies of dabigatran etexilate, which compared bid doses of 110 mg and/or 150 mg with warfarin, but subgroup analyses of bleeding risk are not yet available.[381,382] Product information recommends against the combination of dabigatran etexilate with clopidogrel and other thienopyridines, given alone or as dual antiplatelet therapy with aspirin.

2.4 Antithrombotic Effects: A number of recent phase III clinical trials have evaluated the use of dabigatran etexilate for the prevention and treatment of VTE and to prevent systemic embolism in nonvalvular AF. A phase II study explored its dose-response in acute coronary syndromes.

The efficacy and safety of anticoagulant prophylaxis for VTE depends in any given patient population on the dose, timing, duration, and choice of drug therapy (ie, on the treatment regimen as well as the therapeutic agent). The VTE prevention trials of dabigatran etexilate have evaluated two once-daily dabigatran dosing regimens in patients having elective hip or knee arthroplasty: 150 mg/d or 220 mg/d, beginning soon after surgery with a half-dose of 75 mg or 110 mg. In addition, these two dabigatran regimens were compared with one of two standard enoxaparin prophylaxis regimens: 40 mg/d starting on the evening before surgery, or 30 mg bid starting 12 to 24 h after surgery (worldwide, these are the low-molecular-weight heparin prophylaxis regimens most widely used for joint surgery; both are accepted as standard clinical practice). Two of the trials compared both of these dabigatran regimens with the 40 mg/d schedule of enoxaparin in hip or knee arthroplasty,[383,384] one compared both of the dabigatran regimens with 30 mg enoxaparin bid after knee arthroplasty,[385] and one compared only the 220 mg/d dabigatran etexilate regimen with enoxaparin 40 mg/d in hip arthroplasty.[386]

In all prophylaxis studies, the primary measure of treatment effectiveness was the rate of "total VTE" (a composite of subclinical deep leg vein thrombosis detected with routine bilateral ascending venography done at the end of treatment, confirmed symptomatic DVT or pulmonary embolism (PE), and death from any cause), an outcome that could be evaluated in 73% to 78% of patients. All recorded the presence of major bleeding, clinically relevant nonmajor bleeding, and any adjudicated bleeding.

The two dabigatran prophylaxis regimens were "noninferior" to enoxaparin 40 mg/d in hip arthroplasty and in knee arthroplasty (ie, statistically no less effective),[383,384,387] but each was less effective after knee arthroplasty than 30 mg enoxaparin bid.[385] The first dose of dabigatran etexilate was given 1 to 4 h after surgery in the comparisons with 40 mg/d enoxaparin, but delayed until 12 to 24 h after operation in the comparison with bid enoxaparin. It is likely the bid enoxaparin regimen was superior in part due to the higher daily dose (60 mg), but delaying the start of dabigatran etexilate after surgery may also have contributed to an inferior result. In a subsequent comparison, 220 mg/d dabigatran was again "noninferior" to 40 mg/d enoxaparin in preventing "total VTE" after hip arthroplasty but appeared also to be more effective in preventing "major VTE" (a composite of proximal DVT, nonfatal PE, and death related to VTE).[386] It is now standard practice to start anticoagulant treatment after VTE with a heparin (unfractionated or low molecular weight) together with a VKA, overlap the two drugs for at least 5 days, and stop the heparin only after the INR exceeds 2.0 for > 2 consecutive days.

Dabigatran etexilate was noninferior to warfarin in a double-blind, placebo-controlled and randomized, phase III comparison in patients with deep leg vein thrombosis or PE (the RE-COVER trial) in which study treatment began with at least 5 days of an approved anticoagulant (predominantly unfractionated or low-molecular-weight heparin) plus daily warfarin placebo or warfarin.[382] This study design is unlike that of other (ongoing) evaluations of new oral anticoagulants for VTE, in which the new oral anticoagulant is given alone from the initiation of therapy. The primary measure of efficacy was the incidence of symptomatic and confirmed nonfatal recurrence, or death related to VTE. Randomization in RE-COVER was substratified for a clinical presentation with symptomatic PE (31% of the total) and subgroup analysis suggests this presentation had little or no effect on the relative efficacy of dabigatran. Masking was preserved by using coded POC machines to generate a real or sham INR. Once this exceeded 2.0 for 2 consecutive days, they received dabigatran etexilate 150 mg or its placebo bid (ie, they continued warfarin or started dabigatran). Hence, all patients received at least 5 days of initial treatment with a heparin. However, the median total duration of parenteral therapy before starting dabigatran was 9 days in both of the RE-COVER study groups, which seems longer than the usual clinical practice. INR was within its targeted range of 2.0 to 3.0 for 60% of study time in the warfarin-treated patients.

VKAs are highly effective in preventing embolic stroke from AF, achieving risk reductions of almost 70% when compared with placebo and about 50% when compared with aspirin[387]; their most feared complication is intracranial bleeding, with an added

risk of approximately 0.2% per annum during ongoing warfarin treatment.[388]

Dabigatran dosing regimens of 110 mg and 150 mg bid were compared double-blind with open-label warfarin (target INR, 2.0-3.0) in the RE-LY trial where patients with nonvalvular AF received study treatment of at least 1 year and median of 2 years, and the primary study outcome was ischemic stroke or systemic embolism.[381] On average, the INR in patients taking warfarin was within its target range of 2.0 to 3.0 for 64% of study time. Both efficacy and bleeding risk of dabigatran etexilate depended on the dose. Annual rates of ischemic stroke or embolism were 1.69% with warfarin, 1.53% with 110 mg dabigatran bid (noninferior), and 1.11% with 150 mg dabigatran ($P < .001$ relative to warfarin, $P = .005$ relative to 110 mg dabigatran). The annual incidence of major bleeding was similar, with 150 mg dabigatran and warfarin, but significantly less with 110 mg dabigatran (see section 2.6 "Practical Issues Related to Initiation and Maintenance"). A striking result was the reduced frequency of hemorrhagic stroke with dabigatran compared with warfarin ($P < .001$), regardless of dabigatran dose.

2.5 Monitoring Anticoagulant Intensity

Ex vivo effects on laboratory tests after taking dabigatran etexilate were summarized in section 2.2 "Pharmacokinetics and Pharmacodynamics." Simple widely used coagulation tests have limitations for measuring dabigatran effect: the PT and aPTT are relatively insensitive, the relation between aPTT and dabigatran concentration is nonlinear, and the usual form of TCT is oversensitive. Most promising is the ECT, which has a linear dose-response throughout the range of concentrations expected during prophylaxis or therapy; however, this test is not readily available.[369]

There is no evidence relevant to the possible clinical benefits from laboratory testing, since the phase 3 studies evaluated fixed doses and their reports have not examined clinical outcomes in relation to drug levels or clotting test results. The intent has been to recommend standard doses for most patients; although first principles suggest that laboratory-assisted dose adjustment of this mainly renally excreted drug could add clinical value in selected populations, as in elderly subjects with reduced renal function,[389] it seems unlikely that routine monitoring would yield any wide clinical benefit.[390] Perhaps the most likely role for laboratory testing may be in treated patients who bleed or develop thrombosis, need an acute invasive procedure, or could have taken an overdose. In the setting of major bleeding, or if urgent or emergent surgery is required, a normal TCT rules out the presence of dabigatran. TCT tests are available routinely in many laboratories.

2.6 Practical Issues Related to Initiation and Maintenance

Unlike other new oral anticoagulants, dabigatran etexilate offers a choice between higher and lower dosing regimens, since the schedules evaluated in phase III clinical studies to date were 150 mg and 220 mg once daily when used to prevent postoperative VTE, 150 mg bid for the treatment of VTE, and 110 mg and 150 mg bid in patients with AF.

Dose was important during long-term therapy of patients with AF, in whom 150 mg bid was superior to warfarin and 110 mg bid was noninferior, whereas the rates of major bleeding were similar with 150 mg dabigatran and warfarin but significantly less with 110 mg dabigatran. However, only the 150-mg dose was approved in the United States, whereas in Canada both doses are approved. By contrast, there was less apparent effect of prophylactic dosing level on efficacy or the likelihood of surgical bleeding after major joint replacement, when compared with 40 mg/d enoxaparin.

Subgroup analyses of large clinical studies have provided little evidence to date for clinically important effects of age, gender, or renal function on efficacy or safety within any dosing level. Without such information, the choice between available dosing regimens must rest on individual decisions about the preferred balance between efficacy and bleeding risk and on extrapolations about that risk from pharmacokinetics of this predominantly renally excreted drug. For example, a lower dose may be preferred for elderly patients with a mild to moderate decrease of renal function. In addition, the choice will be influenced by known drug interactions (see section 2.3 "Interactions").

At this time, the only product information approved by European Medicines Agency (EMA) and other regulators is for prophylactic dosing with dabigatran etexilate. Based on clinical study results and pharmacokinetic evidence, EMA recommends a standard daily dose of 220 mg, reduced to 150 mg if patients are aged > 75 years, have a creatinine clearance between 30 and 50 mL/min, or need treatment with verapamil or amiodarone. Ensuring compliance with continued dosing will be a concern for dabigatran etexilate, as it is with any chronic therapy that does not require repeated measurement of drug effect.

2.7 Adverse Events

2.7.1 Bleeding Events: The major potential hazard from dabigatran etexilate is bleeding. After an

elective hip or knee replacement, the risks of major bleeding were small during VTE prophylaxis with dabigatran etexilate and similar to those with prophylactic enoxaparin. The rates were 0.6% to 1.3% in patients given 150 mg/d and 0.6% to 2.0% when the dose was 220 mg/d. These compared with 1.3% to 1.6% with 40 mg/d of enoxaparin and 1.4% with 30 mg of enoxaparin given bid. About 90% of the bleeding events were related to the surgical wound.

Bleeding risk during 6 months of treatment after VTE was less with 150 mg dabigatran bid than with warfarin (target INR, 2.0-3.0). The 6-month rate of "any bleeding" was reduced from 21.9% with warfarin to 16.1% (relative risk, 0.71; 95% CI, 0.59-0.85), and that of "major or clinically relevant but nonmajor" bleeding was diminished from 8.8% to 5.6% (relative risk, 0.63; 95% CI, 0.47-0.84; $P = .002$). Major bleeding was infrequent (1.9% with warfarin and 1.6% with dabigatran).[382]

As indicated earlier, the bleeding risk was dose dependent in the RE-LY study of patients with AF: the annual incidence of major bleeding was 3.4% in patients taking warfarin, 3.1% with 150 mg bid dabigatran (relative risk, 0.93; $P = .31$), and 2.7% if the dose was 110 mg (relative risk, 0.80; 95% CI, 0.69-0.93; $P = .003$); the P value of the difference between dabigatran dosing groups almost reached the conventional level for statistical significance ($P = .052$). Both dabigatran regimens reduced the likelihood of intracranial bleeding by more than one-half compared with warfarin (annual incidence was 0.74% with warfarin, 0.30% with 150 mg dabigatran [$P < .001$] and 0.23% with 110 mg dabigatran [$P < .001$]), whereas annual rates of major extracranial bleeding were little different between the study groups.[381]

Dabigatran etexilate does cause a dose-dependent increase in the incidence of major GI bleeding. This was a statistically significant effect in the RE-LY study; annual rates were 1.02% with warfarin, 1.51% with 150 mg dabigatran ($P < .001$ compared with warfarin), and 1.12% with 110 mg dabigatran.[381] An excess risk was also reported in the RE-COVER trial, in which more patients treated with bid 150 mg dabigatran etexilate had GI bleeding (53 events vs 35 with warfarin).[382]

2.7.2 Nonhemorrhagic Adverse Events: Dyspepsia is a consistently reported adverse effect of dabigatran etexilate. In the RE-LY study, in which warfarin was open label, 11.8% and 11.3% of patients given 110 mg or 150 mg dabigatran complained of dyspepsia compared with 5.8% of patients taking warfarin.[381] Fewer patients reported dyspepsia in the wholly double-blind RE-COVER trial: 2.9% with 150 mg dabigatran bid, and 0.6% with warfarin ($P < .001$).[382]

It is essential to exclude liver toxicity when evaluating new oral anticoagulants, especially because the first orally available direct thrombin inhibitor (ximelagatran) caused unacceptable hepatic toxicity that remains unexplained. About 0.5% of the patients treated with ximelagatran for >35 days developed laboratory evidence of severe liver injury: a greater than threefold elevation of serum alanine transaminase level plus a greater than twofold increase in bilirubin concentration, compared with upper normal limits.[355,391] In one 35-day VTE prevention study, severe hepatic injury with onset after the end of drug exposure was not predicted by regular liver enzyme monitoring during therapy.[392] It has been estimated that ximelagatran may have caused or contributed to fatal liver damage in about 1 in 2,000 patients.[355] No dabigatran study to date has recorded a disproportionate increase of liver enzyme levels or clinically significant liver disease as an adverse effect.

In some ximelagatran studies, short- or long-term treatment was also associated with trends or a statistically significant excess of adverse events related to coronary artery disease.[355] This raised concerns that other new oral anticoagulants might increase the risk from acute cardiac events. Adjudicated coronary events (unstable angina, myocardial infarction, cardiac death) were rare and evenly distributed between study groups in the orthopedic surgery studies of dabigatran etexilate and in the RE-COVER trial, but the RE-LY trial reported an excess of myocardial infarction among the patients treated with dabigatran. The relative risks, compared with warfarin, were 1.35 ($P = .07$) with 110 mg and 1.38 ($P = .048$) with 150 mg dabigatran (the corresponding absolute risks were 0.53%, 0.72%, and 0.74% per annum).[381] There are no ready explanations (although indirect evidence of platelet activation by dabigatran was reported from the PETRO study).[372]

2.8 Reversal of Drug Effect

There is insufficient clinical experience to firmly guide the management of major bleeding, suspected overdose, urgently needed surgery, or urgent invasive diagnostic or therapeutic procedures in patients who are taking this new drug. Pharmacokinetic modeling does, however, indicate how long it takes for drug effects to dissipate after stopping dabigatran etexilate before an elective intervention, although conclusions about the time taken before a return to normal hemostasis remain tentative pending well-documented information about clinical outcomes.

The half-life of dabigatran suggests that drug levels and drug effects should decrease by about 50% at 12 to 18 h after the most recent dose, and the trough levels to 25% of their previous steady state by 24 h

after stopping dabigatran etexilate, so long as creatinine clearance exceeds 50 mL/min. The level at which it is safe to undertake surgery or an invasive procedure is unknown. Moderately severe renal dysfunction (creatinine clearance of 30-50 mL/min) extends the half-life to about 18 h, in which case, or if the surgical bleeding risk is critically high (as with intracranial surgery), it may be better to delay elective procedures until 2 to 4 days after stopping the drug. Measuring the TCT or aPTT should help to estimate the residual level of dabigatran.[369]

In addition to immediately stopping drug administration, the clinical management of major bleeding would require early volume and RBC replacement, urgent assessment for cause, and any local measures that may be required until the bleeding stops (pressure, cautery, suture, or other interventions). It is believed that maintaining an adequate diuresis could help to protect the renal excretion of dabigatran. Although product information for dabigatran etexilate mentions the use of fresh frozen plasma to help control bleeding, this seems unlikely to influence the drug effects. Thus plasma should only be administered in the setting of a documented dilutional coagulopathy.

Dabigatran has no antidote, and the management of life-threatening bleeding remains empirical. Indirect evidence from animal models and in vitro studies suggests that recombinant factor VIIa or a prothrombin complex concentrate may bypass the anticoagulant effects of high dabigatran concentrations.[369] It may also be relevant that hemodialysis removed 62% of circulating dabigatran within 2 h and 68% within 4 h in an open-label study of 12 subjects with end-stage renal failure who received 50 mg dabigatran etexilate.[369,370] In vitro mixing experiments suggest that early administration of activated charcoal might reduce the absorption of dabigatran etexilate.[369]

3.0 Direct Factor Xa Inhibitors: Rivaroxaban

Rivaroxaban is a direct factor Xa inhibitor and is currently approved in many countries, including the United States, for the prevention of VTE in patients undergoing total hip or knee replacement surgery. The drug is undergoing an extensive clinical development program in other clinical settings, including the treatment of VTE and the prevention of acute ischemic stroke in patients with AF. In phase III clinical trials, rivaroxaban was found to be more effective than the low-molecular-weight heparin enoxaparin in preventing VTE after total hip or knee replacement surgery[393-396] and more effective than placebo when given to patients with previous DVT or PE after an initial 6- to 12-month course of standard oral anticoagulant treatment. The doses evaluated were 10 mg once daily in major orthopedic surgery and 20 mg once daily in the long-term secondary prevention of VTE.

3.1 Pharmacology

Rivaroxaban is a selective and competitive active-site-directed, reversible factor Xa inhibitor with selectivity for factor Xa that is > 10,000-fold that for other trypsin-like serine proteases.[397] Rivaroxaban acts through electrostatic interaction with Asp189 in the S1 pocket of factor Xa. This interaction involves the chlorine substituent of the chlorothiophene moiety, which interacts with the aromatic ring of Tyr228 at the bottom of the S1 pocket.[398] Because rivaroxaban is a nonbasic, small molecule (the molecular weight is 436 g/mol) it can inhibit not only free factor Xa but also prothrombinase complex and clot-associated factor Xa.[399,400] This range of activities is unique to small, direct inhibitors because the factor Xa incorporated in the prothrombinase complex is protected from inhibition by antithrombin and by antithrombin-dependent anticoagulants.

Rivaroxaban is minimally soluble in organic solvents and nearly insoluble in water.[398] Binding to plasma proteins, mainly albumin, ranges between 92% and 95%.[398] Rivaroxaban is metabolized via oxidative and hydrolytic pathways involving different classes of enzymes. In humans, CYP3A4 and CYP2J2 are the two enzymes responsible for its oxidative metabolism and contribute to a similar extent.[401] About 66% of rivaroxaban is excreted via the kidneys (36% unchanged), and 28% is excreted in the feces (7% unchanged).

Inhibition of factor Xa activity by rivaroxaban is highly dependent on the concentration of the drug. Rivaroxaban induces prolongation of the PT, aPTT, and heparin clotting time, among other tests.[399,402] On the other hand, rivaroxaban does not affect the bleeding time or platelet aggregation.[399,403] Animal models demonstrated dose-dependent reduction of thrombus formation by rivaroxaban.[399]

3.2 Pharmacokinetics and Pharmacodynamics

In healthy men aged 19 to 45 years, single doses of rivaroxaban administered after a fasting period of 10 h produced a median inhibition of factor Xa activity that ranged from 20% with the 5-mg dose to 61% with the maximum dose of 80 mg.[404] No significant inhibition was observed with doses of < 5 mg. The maximum inhibition of factor Xa activity occurred between 1 and 4 h after drug administration, and the half-life of the biologic effect was 6 to 7 h. Factor Xa activities did not return to normal until after 24 h, when doses > 5 mg were administered.[404] The effect on PT prolongation had a similar profile, as did effects

on the aPTT and HEPTEST (a low-molecular-weight heparin activity assay). Conversely, rivaroxaban had no effect on thrombin and antithrombin activity.[404] Peak plasma concentration of the drug occurred at 2 h, and the terminal half-life was between 7 and 17 h. At doses > 10 mg, the increases in peak plasma concentration and AUC were less than dose proportional. Approximately 40% of the administered dose was excreted unchanged via the kidneys when the 1.25-mg dose was administered; this proportion decreased to 10% with the highest doses tested (ie, 60-80 mg). Finally, inhibition of factor Xa activity and PT prolongation correlated strongly with plasma concentrations ($r = 0.949$ and $r = 0.935$, respectively).[404] When multiple doses were administered at mealtime in healthy male subjects aged 20 to 45 years, the maximum inhibition of factor Xa activity was documented after approximately 3 h and it was dose dependent, ranging from 22% after 5 mg to 68% after 30 mg (Table 4).[405] Inhibition was maintained for 8 to 12 h after 5-mg doses and for approximately 12 h after doses of 10 mg to 30 mg. Daily rivaroxaban doses did not cause a further increase in the maximum inhibition of factor Xa activity. A very similar pattern was observed with PT, aPTT, and HEPTEST prolongation, which was dose dependent for all tests, reached maximum levels after 1 to 4 h, and was comparable after the first and last administered dose.[405] When inhibition of factor Xa activity was compared after once, twice, or three times daily administration of the 5-mg dose, there was no detectable difference between the maximum effect on the first and the last day of administration. The plasma concentrations of rivaroxaban were also dose dependent, with maximum concentrations at 3 to 4 h for all doses and regimens and a half-life of approximately 4 to 6 h on the first day and of approximately 6 to 9 h on the last day of treatment. The correlation between plasma concentrations of rivaroxaban and inhibition of factor Xa activity or PT prolongation was linear, with an r of 0.950 and 0.958, respectively.[405]

Data from this study were later used for population modeling that would predict the behavior of the drug in patients.[406] Results of this model showed that pharmacokinetics of rivaroxaban were linear and dose proportional up to the 30-mg dose and confirmed a strong linear correlation between plasma concentration and pharmacodynamic parameters.

The pharmacokinetics and pharmacodynamics of rivaroxaban were also measured in a population of healthy elderly subjects and in patients with an extremely low or extremely high body weight (Table 4).[407,408] In the first study, healthy subjects aged ≥ 60 years received daily doses ranging between 30 mg and 50 mg after a standard breakfast.[407] It was intended to assess even higher doses of up to 80 mg, but dose escalation was stopped prematurely because of an apparent ceiling effect in pharmacokinetic and pharmacodynamic variables. Maximum inhibition of factor Xa activity ranged from 68.4% after the 30-mg dose to 75.3% and 74.5% after the doses of 40 mg and 50 mg, respectively, and occurred 2 to 4 h after administration. No gender differences were observed. PT prolongation was also greater with the 40-mg dose than with the 30-mg dose, and was not different between the two highest doses, with a maximum effect 2 to 3 h after drug administration. Similar profiles were documented for the aPTT and the HEPTEST. Maximum concentration of rivaroxaban was reached after 4 h in all dose groups, and the half-life

Table 4—*[Section 3.2] Pharmacokinetic Characteristics of Rivaroxaban*[402,404,405]

Population and Dosage	Tmax, h, median	Cmax, μg/L	AUC or AUCτ, μg·h/L	$t_{1/2}$, h
Healthy adults[a]				
5 mg od	3.00	76.4 (18.3)	505.5 (19.7)	8.4 (32.6)
5 mg bid	3.00	85.3 (17.7)	458.5 (13.1)	7.0 (27.8)
5 mg tid	2.00	123.8 (19.7)	557.3 (20.4)	5.8 (35.5)
10 mg bid	2.98	158.0 (18.8)	863.8 (18.6)	7.6 (26.7)
20 mg bid	2.50	318.1 (18.7)	1,903.0 (24.5)	8.0 (40.7)
30 mg bid	3.02	451.9 (10.5)	2,728.0 (14.6)	9.2 (64.1)
Healthy elderly				
30 mg single dose	4.0	392.0 (23.0)	3,531.0 (20.0)	11.7 (63.7)
40 mg single dose	4.0	461.0 (16.8)	4,385.0 (24.1)	13.3 (31.9)
50 mg single dose	4.0	437.0 (32.0)	4,496.0 (33.9)	11.9 (47.8)
Body weight[b]				
≤ 50 kg women only	4.00	178.1 (16.6)	1,172.0 (22.0)	9.6 (36.7)
70-80 kg	3.50	143.4 (26.5)	1,029 (20.1)	7.2 (42.1)
> 120 kg	4.00	149.0 (20.4)	1,155 (15.6)	7.3 (25.4)

Data are presented as geometric mean (geometric coefficient of variation). AUCτ = AUC during a dosage interval at steady state. See Table 3 for expansion of other abbreviations.

[a]Multiple doses.

[b]10 mg single dose.

ranged between 12 and 13 h (Table 4). Excretion via the kidneys was similar among the dosing groups: although women had higher drug concentrations than men after the 30-mg and 40-mg doses, the authors considered that none of the pharmacokinetic parameters showed clinically significant differences between females and males. Finally, the slopes of correlations between plasma concentrations of rivaroxaban and clotting tests were similar in this population of healthy elderly subjects and the population of healthy young men enrolled in a previous study.[407] Other studies have subsequently confirmed the absence of gender-based differences in the pharmacokinetics of rivaroxaban.[409] Pharmacokinetics and pharmacodynamics of rivaroxaban in young and elderly Chinese men and women are comparable to those previously reported in studies of healthy whites.[410,411]

The effects of extreme body weight on a single 10-mg daily dose of rivaroxaban were assessed in a study carried out in men and women aged 18 to 55 years weighing ≤ 50 kg, between 70 and 80 kg, or > 120 kg (Table 4).[408] Peak plasma concentrations were higher in the low body weight group than in the other two groups, with a Cmax value that was significantly increased by 24% as compared with the normal body weight group. On the other hand, there was no difference in plasma concentrations between subjects with a normal or high body weight, and the AUC was similar among all three groups. The half-life of rivaroxaban was increased by 2 h in the low body weight group. Inhibition of factor Xa activity was similar in the three groups, with a maximum effect that occurred after 3 to 4 h and was slightly lower in the group of patients weighing > 120 kg than in the other two groups. The maximum prolongation of PT, aPTT, and the HEPTEST decreased significantly with increasing body weight, although overall all these changes were small.[408]

Finally, to define the pharmacokinetics and pharmacodynamics of rivaroxaban in a population of patients undergoing major orthopedic surgery, blood samples were obtained from patients enrolled in phase II randomized trials.[412,413] The first study analyzed data from two trials that assessed bid doses of rivaroxaban ranging from 2.5 mg to 30 mg administered at mealtimes.[412] Exclusion criteria from these trials were body weight < 45 kg, creatinine clearance < 30 mL/min, and severely impaired hepatic function. The authors found substantial variability in the pharmacokinetics of rivaroxaban on the first postoperative day for all doses, possibly attributed to the presence of "slow" and "fast" absorbers. Overall, absorption was fast, with a Cmax within 1 to 2 h. Absorption status was unrelated to gender, food, anesthesia, or the use of comedications. The variability at steady state was only moderate, and the increase of exposure was dose dependent. Between 14% and 31% of rivaroxaban was excreted unchanged via the kidneys. The clearance of the drug was consistently affected only by renal function, with a creatinine clearance of 30 mL/min resulting in a 15% to 35% higher drug exposure in these two trials.[412] Based on the results of subsequent simulations of various scenarios for patients with different extremes of covariates, it was estimated that plasma concentrations of rivaroxaban would have exceeded the 90% CI of the average patient only in a 90-year-old patient weighing 30 kg, suggesting that one dose of the drug could be administered to all patients regardless of their age, gender, and body weight. The second study compared data from two phase II clinical trials, both carried out in patients undergoing total hip replacement surgery, one assessing bid doses and the other assessing single daily doses.[413] Overall, the pharmacokinetics were similar after once-daily or bid dosing, since only the steady-state area under the plasma concentration-time curve over 24 h was higher with once-daily doses than with bid doses[413] (Table S4). PT prolongation strongly correlated with plasma concentrations when rivaroxaban was given once daily or bid.[413] Simulations of plasma concentration expected after a 10-mg once-daily dose in model subjects with extreme demographic characteristics showed that plasma concentration-time profiles would fall within the predicted 90% CI of the average patient.

3.3 Interactions

3.3.1 Drugs: In humans, CYP3A4 plays a pivotal role in the oxidative metabolism of rivaroxaban.[401] Drugs that inhibit or induce CYP3A4 have the potential to interact with rivaroxaban. However, only drugs that act as strong inhibitors of both CYP3A4 and of P-glycoprotein, a transporter protein of which rivaroxaban is a substrate, have been shown to cause important reduction of the clearance of the drug, thus provoking a significant increase in plasma concentrations. These drugs include azole antimycotics and HIV protease inhibitors. The concomitant administration of rivaroxaban with ketoconazole 400 mg once daily or with ritonavir 600 mg bid resulted in an approximately 2.5-fold increase in the mean AUC and 1.7-fold increase in the mean Cmax of rivaroxaban, together with significantly increased effects on clotting tests.[414] Thus, the use of rivaroxaban in patients receiving ketoconazole, itraconazole, voriconazole, and posaconazole or any HIV protease inhibitor is currently not recommended.[414] Rivaroxaban should be used with caution when given together with other drugs that strongly inhibit only CYP3A4 or only P-glycoprotein. These drugs include clarithromycin, which at the dose of 500 mg bid leads to a

1.5-fold increase in the mean AUC and to a 1.4-fold increase in the Cmax of the drug, and erythromycin, which at the dose of 500 mg three times daily causes a 1.3-fold increase in both the AUC and the Cmax.[414] On the other hand, reduced plasma concentrations of rivaroxaban can occur when strong CYP3A4 inducers are coadministered. These include rifampicin (which causes a 50% decrease in the AUC), phenytoin, carbamazepine, phenobarbital, or St. John's wort.[414] The concomitant administration of rivaroxaban with substrates of either CYP3A4 or P-glycoprotein, such as atorvastatin, digoxin, or midazolam, did not result in clinically relevant interactions.[415,416]

The concomitant administration of rivaroxaban and aspirin was tested in a randomized, two-way crossover study.[403] Healthy men aged between 18 and 55 years were randomized to receive aspirin alone at the dose of 500 mg on the first day and 100 mg on the second day, rivaroxaban alone at the dose of 15 mg, or aspirin and rivaroxaban at the same dosages. Patients receiving rivaroxaban alone were subsequently treated with rivaroxaban and aspirin and vice versa. Maximum levels of inhibition of factor Xa activity and maximum prolongation of the PT, aPTT, and the HEPTEST were similar in patients treated with rivaroxaban alone and in patients treated with rivaroxaban plus aspirin. Collagen-stimulated platelet aggregation was inhibited in the aspirin-alone group and in the aspirin plus rivaroxaban group but not in the rivaroxaban-alone group. Inhibition of platelet aggregation with aspirin was 89.3% greater than with rivaroxaban alone, and with aspirin plus rivaroxaban it was 97.4% greater than with rivaroxaban alone. Bleeding time was not affected by rivaroxaban alone, but was prolonged to 1.46 times the baseline by aspirin alone and to 1.96 times the baseline by adding aspirin to rivaroxaban. The combination of aspirin with rivaroxaban prolonged bleeding time more than aspirin alone. Pharmacokinetic parameters of rivaroxaban were not altered by the coadministration of aspirin.

The concomitant administration of rivaroxaban and a nonsteroidal antiinflammatory drug was also tested in a randomized two-way crossover study.[417] Healthy men aged between 18 and 45 years were randomized to receive 15 mg of rivaroxaban alone or 15 mg rivaroxaban plus naproxen 500 mg. After 14 days all subjects crossed over. Maximum inhibition of factor Xa activity and maximum prolongations of the PT, aPTT, and HEPTEST were similar between the two groups. Bleeding time was significantly increased by the combined use of rivaroxaban and naproxen as compared with rivaroxaban alone. Plasma concentrations of rivaroxaban were slightly increased, by 10% for both the AUC and the Cmax after the coadministration of naproxen. Finally, the concomitant administration of drugs that alter the gastric pH, such as the histamine H2 antagonist ranitidine or the antacid aluminum-magnesium hydroxide, had no effect on the plasma concentrations or pharmacodynamics of rivaroxaban.[418]

3.3.3 Environmental Factors: The effect of food on the absorption of rivaroxaban was tested in healthy male subjects aged 18 to 45 years.[418] After a single dose of 5 mg, the absorption of rivaroxaban was found to be slower in the fed state than the fasting state, with a delay in the median time to reach the peak plasma concentration from 2.75 h to 4.0 h. The AUC and the Cmax both increased with the concomitant administration of food, by 28% and 41%, whereas the terminal half-life remained unchanged. The fed state increased the time to maximum prolongation of the PT and maximum inhibition of factor Xa activity and also the maximum effect on these clotting tests. Thus, absorption of rivaroxaban was moderately increased after the administration of food, with a resulting increase in pharmacokinetic and pharmacodynamic parameters. In addition, concomitant food intake reduced interindividual variability, whereas elimination remained unchanged. No pharmacokinetic differences were documented when high-fat, high-calorie meals were compared with high-carbohydrate meals. These food effects have been attributed to a prolonged length of stay in the stomach that is possibly related to the lipophilicity and limited aqueous solubility of the drug.[418]

3.4 Antithrombotic Effect

Rivaroxaban acts as a competitive inhibitor of the amidolytic activity of factor Xa. In vitro, the inhibitory effect is concentration dependent, with an inhibitory constant against factor Xa of 0.4 ± 0.02 nM.[399] The selectivity of rivaroxaban for factor Xa is 10,000-fold greater than for other serine proteases, including factor Va, factor IXa, factor XIa, thrombin, and activated protein C. The onset of action is rapid and reversible, as shown by the kinetic association rate constant of 1.7×10^7 mol/L^{-1}/s^{-1} and kinetic dissociation rate constant of 5×10^{-3}/s^{-1}.[419] The inhibitory effect is maintained when factor Xa is complexed with factor Va and Ca^{2+} on a phospholipid membrane (prothrombinase bound) as shown by a concentration-dependent inhibition of thrombin generation[399] and also when factor Xa is bound to the clot.[420]

In plasma, endogenous human factor Xa is inhibited by rivaroxaban to cause dose-dependent prolongation of both the PT and aPTT, with the PT being more sensitive than the aPTT.[399] In whole blood and in platelet-rich plasma, rivaroxaban has a dose-dependent effect in prolonging the initiation phase of thrombin generation after activation of the tissue

factor pathway, reduces the maximum concentration of generated thrombin, and decreases the endogenous thrombin potential.[400] The effect on the initiation and propagation phases of thrombin generation is greater than the effect on the decay phase, expressed by the endogenous thrombin potential. High concentrations of rivaroxaban are able to prevent thrombin generation almost completely, due to inhibition of factor Xa bound to the prothrombinase complex.[400] The effect of rivaroxaban on platelet-induced thrombin generation in vivo was tested in a randomized, placebo-controlled, crossover study of 12 healthy male subjects aged 27 to 37 years.[397] The prothrombinase-induced clotting time, a plasma clotting assay based on the activation of coagulation using factor Xa, phospholipids, and an enzyme that activates factor V, was prolonged dose dependently by rivaroxaban, with a maximum effect after 2 h. Thrombin generation in platelet-rich plasma was markedly reduced by rivaroxaban, with an 80% to 90% decrease (by 5-mg and 30-mg doses, respectively) of the peak collagen-induced endogenous thrombin potential at 2 h. There was a close correlation between plasma concentration of rivaroxaban and its effects on factor Xa activity inhibition, prothrombinase-induced clotting time prolongation, and endogenous thrombin potential reduction.[397]

Rivaroxaban has no effects on platelet aggregation induced by collagen, adenosine diphosphate, or thrombin.[403,421,422] An indirect effect on platelet aggregation induced by tissue factor, determined by the inhibition of thrombin generation, has been reported in defibrinated plasma.[423]

In venous thrombosis models and in arteriovenous shunt models, rivaroxaban was shown to cause dose-dependent reduction of thrombus formation, inhibition of factor Xa activity, and prolongation of the PT.[399] The antithrombotic effective doses of rivaroxaban did not prolong bleeding time in animal bleeding models.[399]

3.5 Monitoring Anticoagulant Intensity

The predictable pharmacologic profile of rivaroxaban allows the administration of the drug at fixed doses without the need for routine laboratory monitoring or dose adjustments. However, there may be rare situations, as in the case of overdose or unexpected bleeding, assessment of compliance, evaluation of drug interactions, or assessment of drug accumulation in renal or hepatic impairment, when the availability of a quantitative clotting assay might be valuable. Despite the predictable, dose-dependent effects of rivaroxaban on the PT and (to a lesser extent) the aPTT, and on tests measuring thrombin generation, there are currently no validated laboratory assays that can be recommended to monitor rivaroxaban or any recommendations for dose adjustments based on observed test results. For instance, the thromboplastins used for PT clotting assays have differing sensitivities to factor Xa inhibitors, and the INR introduced to correct for differences in PT sensitivity when monitoring the VKAs does not adequately correct for differences in assay sensitivity to direct factor Xa inhibitors. Smith and Morissey[424] evaluated the effects of five commercial thromboplastin reagents on sensitivity of the PT to rivaroxaban and its correlation with the INR. PT ratios (ie, PT with drug/PT without drug) were measured using normal human plasma to which rivaroxaban 1 μg/mL was added in vitro. PT ratios varied from 2.25 to 7.32 with the different thromboplastins; subsequent conversion to an INR further exacerbated the observed differences in sensitivities to rivaroxaban between the various PT assays.

Recently, Samama et al[402] carried out a study that aimed to identify a clotting test suitable for monitoring rivaroxaban activity by evaluating the effects on a number of different assays of increasing the drug concentration in pooled citrated normal human platelet-poor plasma. There was a concentration-dependent prolongation of the PT and aPTT, but the increases in clotting times varied depending on the thromboplastin reagent used.[402] The effect of rivaroxaban on the aPTT was weaker than that on the PT. Rivaroxaban also prolonged the thromboelastograph parameters. Standard methods for the HEPTEST and the prothrombinase-induced clotting time resulted in paradoxical responses. Tests used to measure anti-factor Xa activity of the low-molecular-weight heparins showed dose-dependent effects but were associated with some degrees of variation. Finally, rivaroxaban also caused a dose-dependent increase of the diluted Russell viper venom ratio. Specific calibration of some of these tests may lead to the availability of an appropriate assay to monitor the pharmacodynamic effects of the drug. Neither the PT (expressed either in seconds or as a ratio) nor the aPTT should be used to monitor the anticoagulant effect of rivaroxaban.

3.6 Practical Issues Related to Initiation and Maintenance

Rivaroxaban is currently approved in many countries for the prevention of VTE in patients undergoing total hip replacement surgery and total knee replacement surgery based on the results of four phase 3 clinical trials.[393-396] For this indication, the approved dose is 10 mg once daily. Treatment should be started between 6 and 10 h after surgery, and the duration of treatment should vary from 2 weeks in patients undergoing total knee replacement surgery

to 5 weeks in patients undergoing total hip replacement surgery, although approved durations may vary among countries. Rivaroxaban is currently not approved for use in patients with severe renal failure and a creatinine clearance of < 15 mL/min, in patients with hepatic disease associated with coagulopathy, and in patients receiving concomitant systemic treatment with azole-antimycotics or HIV protease inhibitors.[414] Rivaroxaban is also not approved for use in children or adolescents < 18 years of age because of the absence of clinical data, in pregnant women because of the potential for reproductive toxicity (observed in animals) and the evidence that the drug passes the placenta, and during breast-feeding because the drug is secreted into milk.[414]

Future additional indications are expected following the results of recently published clinical trials. In particular, more information on the long-term administration of rivaroxaban now originates from studies carried out in patients with VTE or AF in whom the drug was administered for a minimum of 6 months and up to a maximum of 40 months.[425,426]

To date, the results of a phase III clinical trials assessing the use of rivaroxaban in the treatment of DVT have been published. In the EINSTEIN study, > 3,400 patients with acute symptomatic proximal DVT were allocated to rivaroxaban 15 mg bid for 3 weeks followed by 20 mg once daily or initial therapy with enoxaparin (1 mg/kg bid) and simultaneous warfarin administered to a target INR of 2.0 to 3.0.[425] The primary outcome measure was symptomatic recurrent VTE, which occurred in 36 rivaroxaban-treated patients and 51 low-molecular-weight heparin/warfarin-treated patients (HR, 0.68; 95% CI, 0.44-1.04; $P < .0001$ for noninferiority). Patients who had completed 6 to 12 months of anticoagulant treatment with either a VKA or rivaroxaban (if also enrolled in the acute-phase rivaroxaban treatment study) after an acute episode of VTE were randomized to receive rivaroxaban 20 mg once daily or matching placebo for additional 6 to 12 months. The study showed superior efficacy of rivaroxaban over placebo in the prevention of recurrent venous thromboembolic events, a nonsignificant increase in major bleeding, and an increase in clinically relevant, nonmajor bleeding events.[425]

More recently, the results of the Rivaroxaban Once Daily Oral Direct Factor Xa inhibition Compared with Vitamin K Antagonist for the Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET) study comparing rivaroxaban 20 mg once daily (15 mg once daily if patients had a creatinine clearance between 30 and 49 mL/min) with warfarin in patients with AF and at least two additional risk factors for embolic events were published. In this randomized, double-blind, controlled trial, rivaroxaban was noninferior to warfarin in the prevention of ischemic stroke or systemic embolism, with a similar rate of major bleeding events but with fewer intracranial and fatal bleeding events.[426]

Other recently completed phase III trials include a study on the prevention of VTE in medical patients (Multicenter, Randomized, Parallel Group Efficacy and Safety Study for the Prevention of VTE in Hospitalized Medically Ill Patients Comparing Rivaroxaban with Enoxaparin [MAGELLAN] study), in which rivaroxaban is administered at the same 10 mg once-daily dose used in the Regulation of Coagulation in Major Orthopedic Surgery Reducing the Risk of DVT and PE (RECORD) program; and the EINSTEIN PE study, in which patients with hemodynamically stable PE are treated for the first 3 weeks with a 15 mg bid dose of rivaroxaban followed by a 20 mg once-daily dose. Finally, a phase III study in patients with acute coronary syndromes is assessing rivaroxaban doses of 2.5 mg and 5.0 mg bid.[427]

3.7 Adverse Events

3.7.1 Bleeding Events: Data on the incidence of bleeding complications with the use of rivaroxaban are currently available from the results of phase II and phase III clinical trials only. As of yet, no published phase IV studies or reports from clinical practice have been published. Reported rates must take into account the definition of bleeding used in each study, given the variability of definitions used among clinical trials assessing different compounds.

The reported rates of any bleeding with rivaroxaban in the four phase III clinical trials carried out in patients undergoing total hip and total knee replacement surgery range from 4.9% to 10.5%[393-396] and are comparable to the rates observed in patients treated with enoxaparin. Rates of major bleeding events vary from < 0.1% to 0.7%, and the rates of clinically relevant nonmajor bleeding events vary from 2.6% to 3.3%, again with no statistically significant differences between the rates observed in the rivaroxaban groups and the rates observed in the comparator groups (ie, the low-molecular-weight heparin enoxaparin in all studies).[393-396] A pooled analysis of the four RECORD studies reported that the concomitant use of nonsteroidal antiinflammatory drugs, aspirin, or other antiplatelet agents was not associated with an increased rate of bleeding events.[428]

The rate of any bleeding event in patients with a previous episode of VTE treated with rivaroxaban 20 mg once daily for 6 to 12 months after the initial standard course of anticoagulant therapy was 23%; major bleeding and clinically relevant nonmajor bleeding occurred in 0.7% and 5.4% of patients, respectively.[425]

3.7.2 Nonhemorrhagic Adverse Events: None of the four phase III clinical trials found evidence of drug-associated liver toxicity in patients treated with rivaroxaban for up to 5 weeks.[393-396] Overall, the proportion of patients with elevated liver enzymes was low in all studies and was similar between rivaroxaban and enoxaparin. The incidence of cardiovascular events was low while on treatment with rivaroxaban, ranging between 0.1% and 0.7%, and was similar to that seen with enoxaparin. Drug-related adverse events occurred in 12% to 20% of patients in the four RECORD trials; these rates were similar to those observed with enoxaparin. The most frequent adverse events were nausea, vomiting, and constipation.

3.8 Reversal

There is currently no specific antidote available to antagonize the effects of rivaroxaban. In case of overdose, the use of activated charcoal to reduce absorption is suggested.[414] Because of the high plasma protein binding, rivaroxaban is unlikely to be dialyzable. In case of active bleeding, possible strategies currently include discontinuation of treatment and administration of blood products or component transfusion if required to treat an identified deficiency.[414] However, there is currently no direct evidence in humans to support the efficacy of blood product transfusion or other interventions in improving hemostasis when patients have received rivaroxaban. Recently, Perzborn et al[429] reported the results of a study carried out in rats treated with high-dose rivaroxaban, which aimed to assess the efficacy of prothrombin complex concentrate. After determination of baseline mesenteric bleeding time, rats were initially treated with IV rivaroxaban and subsequently received four-factor prothrombin complex concentrates at 25 U/kg or 50 U/kg. Prolongation of the bleeding time was almost completely abrogated by higher dose of prothrombin complex concentrates, whereas the lower dose was ineffective. The use of recombinant factor VIIa is also suggested in the presence of life-threatening bleeding based on some preclinical data.[414] A reconstructed recombinant factor Xa has been recently proposed as a potential antidote for factor Xa inhibitors.[430] This is a catalytically inactive factor Xa that has no procoagulant or anticoagulant activity and does not interfere with the prothrombinase complex but maintains high affinity for factor Xa inhibitors. In plasma, the addition of the antidote dose-dependently reversed factor Xa inhibition as measured by anti-factor Xa units, tissue factor-initiated thrombin generation, and clotting assays. In vivo, the antidote completely reversed PT prolongation induced by intravenous infusion of rivaroxaban in rats.

Conclusion

Over the last decades, a large amount of research has been addressed to improve the understanding of the mechanisms of warfarin, acenocoumarol, and phenprocoumon and to improve the management of patients treated with these VKAs. Several studies have in particular identified some genetic factors associated with the individual responses to VKAs and several drugs, foods, and environmental factors that can interact with these compounds. Several induction and maintenance strategies have been compared, and management studies have evaluated different approaches for the monitoring of patients on VKAs, including AMS and AC, computer programs, and POC for INR testing. The results of all such studies have greatly contributed to improve the efficacy and safety of oral anticoagulant therapy and to increase the number of patients who can be deemed eligible for such treatment. Finally, a number of trials have also addressed the management of patients on VKA treatment who are at increased risk of bleeding or are actively bleeding, and a number of therapeutic strategies have been proposed, although additional research may be warranted in particular to further improve the management of the bleeding patient.

The new oral anticoagulant drugs have the potential to overcome several drawbacks of the VKAs. These drugs can be administered at fixed doses and do not require laboratory monitoring, thus offering a clear advantage over the VKAs. It is hoped that new studies will provide us with further information on the role of specific laboratory tests for the monitoring of the activity of these new classes of drugs, when requested, and on the optimal management of drug-related adverse events.

Acknowledgments

Author contributions: As Topic Editor, Dr Ageno oversaw the development of this article, including any analysis and subsequent development of the information contained herein.
Dr Ageno: contributed as Topic Editor.
Dr Gallus: contributed as a panelist.
Dr Wittkowsky: contributed as a panelist.
Dr Crowther: contributed as a panelist.
Dr Hylek: contributed as a panelist.
Dr Palareti: contributed as a panelist.
Financial/nonfinancial disclosures: In summary, the authors have reported to *CHEST* the following conflicts of interest: Dr Ageno has received fees for steering committee or safety advisory boards membership from Bayer Health Care, sanofi-aventis, and Thrombogenics; research grants from GlaxoSmithKline, Alexion Pharmaceutical, Pfizer, and Bayer Health Care; honoraria for speaking activities at national or international meetings and for advisory boards from Bayer Health Care, GlaxoSmithKline, sanofi-aventis, Pfizer, BMS, and Boehringer Ingelheim. Dr Gallus has received fees for steering committee membership from BMS, Pfizer, sanofi-aventis, Bayer, ASTELLAS, Daiichi Sankyo. Dr Crowther has served on various advisory boards, has assisted in the preparation of educational materials, has sat on data and safety monitoring boards, and his institution has received research funds from the following companies: Leo Pharma, Pfizer, Boehringer Ingelheim,

Bayer, Octapharm, CSL Behring, and Artisan. His personal total compensation for these activities over the last 3 years totals less than $10,000. Dr Hylek has served on advisory boards (Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Daiichi Sankyo, Johnson & Johnson, Merck, Ortho-McNeil, and Pfizer; total dollar amount less than $10,000). She has served on the Executive Steering Committee-ARISTOTLE trial sponsored by Bristol-Myers Squibb and Pfizer (dollar amount less than $10,000). She has also served on the Executive Steering Committee-ORBIT AF Registry sponsored by Ortho-McNeil (dollar amount less than $10,000). She has participated in a symposium sponsored by Boehringer Ingelheim (dollar amount less than $10,000). Drs Wittkowsky and Palareti have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of sponsors: The sponsors played no role in the development of these guidelines. Sponsoring organizations cannot recommend panelists or topics, nor are they allowed prepublication access to the manuscripts and recommendations. Guideline panel members, including the chair, and members of the Health & Science Policy Committee are blinded to the funding sources. Further details on the Conflict of Interest Policy are available online at http://chestnet.org.

Endorsements: This guideline is endorsed by the American Association for Clinical Chemistry, the American College of Clinical Pharmacy, the American Society of Health-System Pharmacists, the American Society of Hematology, and the International Society of Thrombosis and Hematosis.

Additional information: The supplement Tables can be found in the Online Data Supplement at http://chestjournal.chestpubs.org/content/141/2_suppl/e44S/suppl/DC1.

References

1. Whitlon DS, Sadowski JA, Suttie JW. Mechanism of coumarin action: significance of vitamin K epoxide reductase inhibition. *Biochemistry*. 1978;17(8):1371-1377.
2. Fasco MJ, Hildebrandt EF, Suttie JW. Evidence that warfarin anticoagulant action involves two distinct reductase activities. *J Biol Chem*. 1982;257(19):11210-11212.
3. Choonara IA, Malia RG, Haynes BP, et al. The relationship between inhibition of vitamin K1 2,3-epoxide reductase and reduction of clotting factor activity with warfarin. *Br J Clin Pharmacol*. 1988;25(1):1-7.
4. Trivedi LS, Rhee M, Galivan JH, Fasco MJ. Normal and warfarin-resistant rat hepatocyte metabolism of vitamin K 2,3-epoxide: evidence for multiple pathways of hydroxyvitamin K formation. *Arch Biochem Biophys*. 1988;264(1):67-73.
5. Stenflo J, Fernlund P, Egan W, Roepstorff P. Vitamin K dependent modifications of glutamic acid residues in prothrombin. *Proc Natl Acad Sci U S A*. 1974;71(7):2730-2733.
6. Nelsestuen GL, Zytkovicz TH, Howard JB. The mode of action of vitamin K. Identification of gamma-carboxyglutamic acid as a component of prothrombin. *J Biol Chem*. 1974; 249(19):6347-6350.
7. Stafford DW. The vitamin K cycle. *J Thromb Haemost*. 2005;3(8):1873-1878.
8. Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G; American College of Chest Physicians. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest*. 2008; 133(suppl 6):160S-198S.
9. Friedman PA, Rosenberg RD, Hauschka PV, Fitz-James A. A spectrum of partially carboxylated prothrombins in the plasmas of coumarin-treated patients. *Biochim Biophys Acta*. 1977;494(1):271-276.
10. Malhotra OP, Nesheim ME, Mann KG. The kinetics of activation of normal and gamma-carboxyglutamic acid-deficient prothrombins. *J Biol Chem*. 1985;260(1):279-287.
11. Nelsestuen GL, Zytkovicz TH, Howard JB. Role of gamma-carboxyglutamic acid. An unusual protein transition required for the calcium-dependent binding of prothrombin to phospholipid. *J Biol Chem*. 1976;251(18):5648-5656.
12. Prendergast FG, Mann KG. Differentiation of metal ion-induced transitions of prothrombin fragment 1. *J Biol Chem*. 1977;252(3):840-850.
13. Borowski M, Furie BC, Bauminger S, Furie B. Prothrombin requires two sequential metal-dependent conformational transitions to bind phospholipid. Conformation-specific antibodies directed against the phospholipid-binding site on prothrombin. *J Biol Chem*. 1986;261(32):14969-14975.
14. Becker R. The importance of factor Xa regulatory pathways in vascular thromboresistance: focus on protein Z. *J Thromb Thrombolysis*. 2005;19(2):135-137.
15. Breckenridge A. Oral anticoagulant drugs: pharmacokinetic aspects. *Semin Hematol*. 1978;15(1):19-26.
16. O'Reilly RA. Vitamin K and the oral anticoagulant drugs. *Annu Rev Med*. 1976;27:245-261.
17. Kelly JG, O'Malley K. Clinical pharmacokinetics of oral anticoagulants. *Clin Pharmacokinet*. 1979;4(1):1-15.
18. O'Reilly R, ed. *Warfarin Metabolism and Drug-Drug Interactions*. New York, NY: Plenum; 1986.
19. Miners JO, Birkett DJ. Cytochrome P4502C9: an enzyme of major importance in human drug metabolism. *Br J Clin Pharmacol*. 1998;45(6):525-538.
20. Godbillon J, Richard J, Gerardin A, Meinertz T, Kasper W, Jähnchen E. Pharmacokinetics of the enantiomers of acenocoumarol in man. *Br J Clin Pharmacol*. 1981;12(5):621-629.
21. Haustein KO. Pharmacokinetic and pharmacodynamic properties of oral anticoagulants, especially phenprocoumon. *Semin Thromb Hemost*. 1999;25(1):5-11.
22. Mentré F, Pousset F, Comets E, et al. Population pharmacokinetic-pharmacodynamic analysis of fluindione in patients. *Clin Pharmacol Ther*. 1998;63(1):64-78.
23. Home page of the Human Cytochrome P450 (CYP) Allele Nomenclature Committee. Web site. http://www.CYPalleles.ki.se. Accessed August 31, 2010.
24. Scordo MG, Pengo V, Spina E, Dahl ML, Gusella M, Padrini R. Influence of CYP2C9 and CYP2C19 genetic polymorphisms on warfarin maintenance dose and metabolic clearance. *Clin Pharmacol Ther*. 2002;72(6):702-710.
25. Wittkowsky AK. Pharmacology of warfarin and related anticoagulants. In: Ansell J, Oertel L, Wittkowsky A, eds. *Managing Oral Anticoagulation Therapy: Clinical and Operational Guidelines*. Vol. 1. St. Louis, MO: Facts and Comparisons; 2003;29(1)-29:15.
26. Loebstein R, Yonath H, Peleg D, et al. Individual variability in sensitivity to warfarin: nature or nurture. *Clin Pharmacol Ther*. 2001;70(2):159-164.
27. Lindh JD, Holm L, Andersson ML, Rane A. Influence of CYP2C9 genotype on warfarin dose requirements—a systematic review and meta-analysis. *Eur J Clin Pharmacol*. 2009;65(4):365-375.
28. Aithal GP, Day CP, Kesteven PJ, Daly AK. Association of polymorphisms in the cytochrome P450 CYP2C9 with warfarin dose requirement and risk of bleeding complications. *Lancet*. 1999;353(9154):717-719.
29. Higashi M, Veenstra DL, Wittkowsky AK, et al. Influence of CYP2C9 genetic variants on the risk of over anticoagulation and of bleeding events during warfarin therapy. *JAMA*. 2002;287(13):1690-1698.
30. Veenstra DL, Blough DK, Higashi MK, et al. CYP2C9 haplotype structure in European American warfarin patients and association with clinical outcomes. *Clin Pharmacol Ther*. 2005;77(5):353-364.

31. Chern HD, Ueng TH, Fu YP, Cheng CW. CYP2C9 polymorphism and warfarin sensitivity in Taiwan Chinese. *Clin Chim Acta*. 2006;367(1-2):108-113.
32. Herman D, Peternel P, Stegnar M, Breskvar K, Dolzan V. A novel sequence variant in exon 7 of CYP2C9 gene (CYP2C9*24) in a patient on warfarin therapy. *Thromb Haemost*. 2006;95(1):192-194.
33. Visser LE, van Schaik RHN, van Vliet M, et al. The risk of bleeding complications in patients with cytochrome P450 CYP2C9*2 or CYP2C9*3 alleles on acenocoumarol or phenprocoumon. *Thromb Haemost*. 2004;92(1):61-66.
34. Takahashi H, Wilkinson GR, Padrini R, Echizen H. CYP2C9 and oral anticoagulation therapy with acenocoumarol and warfarin: similarities yet differences. *Clin Pharmacol Ther*. 2004;75(5):376-380.
35. Beinema M, Brouwers JR, Schalekamp T, Wilffert B. Pharmacogenetic differences between warfarin, acenocoumarol and phenprocoumon. *Thromb Haemost*. 2008;100(6):1052-1057.
36. Zimmermann A, Matschiner JT. Biochemical basis of hereditary resistance to warfarin in the rat. *Biochem Pharmacol*. 1974;23(6):1033-1040.
37. Rost S, Fregin A, Ivaskevicius V, et al. Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature*. 2004;427(6974):537-541.
38. Li T, Chang CY, Jin DY, Lin PJ, Khvorova A, Stafford DW. Identification of the gene for vitamin K epoxide reductase. *Nature*. 2004;427(6974):541-544.
39. D'Andrea G, D'Ambrosio RL, Di Perna P, et al. A polymorphism in the VKORC1 gene is associated with an interindividual variability in the dose-anticoagulant effect of warfarin. *Blood*. 2005;105(2):645-649.
40. Rieder MJ, Reiner AP, Gage BF, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *N Engl J Med*. 2005;352(22):2285-2293.
41. Geisen C, Watzka M, Sittinger K, et al. VKORC1 haplotypes and their impact on the inter-individual and inter-ethnical variability of oral anticoagulation. *Thromb Haemost*. 2005;94(4):773-779.
42. Sconce EA, Khan TI, Wynne HA, et al. The impact of CYP2C9 and VKORC1 genetic polymorphism and patient characteristics upon warfarin dose requirements: proposal for a new dosing regimen. *Blood*. 2005;106(7):2329-2333.
43. Harrington DJ, Underwood S, Morse C, Shearer MJ, Tuddenham EGD, Mumford AD. Pharmacodynamic resistance to warfarin associated with a Val66Met substitution in vitamin K epoxide reductase complex subunit 1. *Thromb Haemost*. 2005;93(1):23-26.
44. Anthony M, Romero K, Malone DC, Hines LE, Higgins L, Woosley RL. Warfarin interactions with substances listed in drug information compendia and in the FDA-approved label for warfarin sodium. *Clin Pharmacol Ther*. 2009;86(4):425-429.
45. Marsh S, King CR, Porche-Sorbet RM, Scott-Horton TJ, Eby CS. Population variation in VKORC1 haplotype structure. *J Thromb Haemost*. 2006;4(2):473-474.
46. Quteineh L, Verstuyft C, Descot C, et al. Vitamin K epoxide reductase (VKORC1) genetic polymorphism is associated to oral anticoagulant overdose. *Thromb Haemost*. 2005;94(3):690-691.
47. Coumadin (Warfarin) [package insert]. Princeton NJ: Bristol Myers Squibb; 2007.
48. Holbrook AM, Pereira JA, Labiris R, et al. Systematic overview of warfarin and its drug and food interactions. *Arch Intern Med*. 2005;165(10):1095-1106.
49. Orme M, Breckenridge A. Enantiomers of warfarin and phenobarbital. *N Engl J Med*. 1976;295(26):1482-1483.
50. Breckenridge A, Orme M, Wesseling H, Lewis RJ, Gibbons R. Pharmacokinetics and pharmacodynamics of the enantiomers of warfarin in man. *Clin Pharmacol Ther*. 1974;15(4):424-430.
51. O'Reilly RA. Studies on the optical enantiomorphs of warfarin in man. *Clin Pharmacol Ther*. 1974;16(2):348-354.
52. O'Reilly RA, Trager WF, Motley CH, Howald W. Stereoselective interaction of phenylbutazone with [$^{12}C/^{13}C$] warfarin pseudoracemates in man. *J Clin Invest*. 1980;65(3):746-753.
53. Toon S, Low LK, Gibaldi M, et al. The warfarin-sulfinpyrazone interaction: stereochemical considerations. *Clin Pharmacol Ther*. 1986;39(1):15-24.
54. O'Reilly RA. The stereoselective interaction of warfarin and metronidazole in man. *N Engl J Med*. 1976;295(7):354-357.
55. O'Reilly RA. Stereoselective interaction of trimethoprim-sulfamethoxazole with the separated enantiomorphs of racemic warfarin in man. *N Engl J Med*. 1980;302(1):33-35.
56. Lewis RJ, Trager WF, Chan KK, et al. Warfarin. Stereochemical aspects of its metabolism and the interaction with phenylbutazone. *J Clin Invest*. 1974;53(6):1607-1617.
57. O'Reilly RA, Trager WF, Rettie AE, Goulart DA. Interaction of amiodarone with racemic warfarin and its separated enantiomorphs in humans. *Clin Pharmacol Ther*. 1987;42(3):290-294.
58. Cropp JS, Bussey HI. A review of enzyme induction of warfarin metabolism with recommendations for patient management. *Pharmacotherapy*. 1997;17(5):917-928.
59. Ng HJ, Crowther MA. Azathioprine and inhibition of the anticoagulant effect of warfarin: evidence from a case report and a literature review. *Am J Geriatr Pharmacother*. 2006;4(1):75-77.
60. O'Reilly RA. Lack of effect of fortified wine ingested during fasting and anticoagulant therapy. *Arch Intern Med*. 1981;141(4):458-459.
61. Bechtold H, Andrassy K, Jähnchen E, et al. Evidence for impaired hepatic vitamin K1 metabolism in patients treated with N-methyl-thiotetrazole cephalosporins. *Thromb Haemost*. 1984;51(3):358-361.
62. Weitekamp MR, Aber RC. Prolonged bleeding times and bleeding diathesis associated with moxalactam administration. *JAMA*. 1983;249(1):69-71.
63. Owens JC, Neely WB, Owen WR. Effect of sodium dextrothyroxine in patients receiving anticoagulants. *N Engl J Med*. 1962;266:76-79.
64. O'Reilly RA, Sahud MA, Robinson AJ. Studies on the interaction of warfarin and clofibrate in man. *Thromb Diath Haemorrh*. 1972;27(2):309-318.
65. Rothschild BM. Hematologic perturbations associated with salicylate. *Clin Pharmacol Ther*. 1979;26(2):145-152.
66. Hylek EM, Heiman H, Skates SJ, Sheehan MA, Singer DE. Acetaminophen and other risk factors for excessive warfarin anticoagulation. *JAMA*. 1998;279(9):657-662.
67. Bell WR. Acetaminophen and warfarin: undesirable synergy. *JAMA*. 1998;279(9):702-703.
68. Mahé I, Bertrand N, Drouet L, et al. Interaction between paracetamol and warfarin in patients: a double-blind, placebo-controlled, randomized study. *Haematologica*. 2006;91(12):1621-1627.
69. Thijssen HH, Soute BA, Vervoort LM, Claessens JG. Paracetamol (acetaminophen) warfarin interaction: NAPQI, the toxic metabolite of paracetamol, is an inhibitor of enzymes in the vitamin K cycle. *Thromb Haemost*. 2004;92(4):797-802.
70. Hylek EM. Paracetamol (acetaminophen) and warfarin interaction: unraveling the pivotal role of the vitamin K cycle. *Thromb Haemost*. 2004;92(4):672-673.

71. Weibert RT, Lorentz SM, Townsend RJ, Cook CE, Klauber MR, Jagger PI. Effect of erythromycin in patients receiving long-term warfarin therapy. *Clin Pharm*. 1989; 8(3):210-214.
72. Lorentz SM, Weibert RT. Potentiation of warfarin anticoagulation by topical testosterone ointment. *Clin Pharmacol Ther*. 1985;4(3):332-334.
73. Udall JA. Human sources and absorption of vitamin K in relation to anticoagulation stability. *JAMA*. 1965;194(2): 127-129.
74. Dale J, Myhre E, Loew D. Bleeding during acetylsalicylic acid and anticoagulant therapy in patients with reduced platelet reactivity after aortic valve replacement. *Am Heart J*. 1980;99(6):746-752.
75. Schulman S, Henriksson K. Interaction of ibuprofen and warfarin on primary haemostasis. *Br J Rheumatol*. 1989; 28(1):46-49.
76. Battistella M, Mamdami MM, Juurlink DN, Rabeneck L, Laupacis A. Risk of upper gastrointestinal hemorrhage in warfarin users treated with nonselective NSAIDs or COX-2 inhibitors. *Arch Intern Med*. 2005;165(2):189-192.
77. Cazenave JP, Packham MA, Guccione MA, Mustard JF. Effects of penicillin G on platelet aggregation, release, and adherence to collagen. *Proc Soc Exp Biol Med*. 1973;142(1): 159-166.
78. Brown CH III, Natelson EA, Bradshaw W, Williams TW Jr, Alfrey CP Jr. The hemostatic defect produced by carbenicillin. *N Engl J Med*. 1974;291(6):265-270.
79. Roth GJ, Majerus PW. The mechanism of the effect of aspirin on human platelets. I. Acetylation of a particulate fraction protein. *J Clin Invest*. 1975;56(3):624-632.
80. Andreotti F, Testa L, Biondi-Zoccai GGL, Crea F. Aspirin plus warfarin compared to aspirin alone after acute coronary syndromes: an updated and comprehensive meta-analysis of 25,307 patients. *Eur Heart J*. 2006;27(5):519-526.
81. Chesebro JH, Fuster V, Elveback LR, et al. Trial of combined warfarin plus dipyridamole or aspirin therapy in prosthetic heart valve replacement: danger of aspirin compared with dipyridamole. *Am J Cardiol*. 1983;51(9):1537-1541.
82. Turpie AGG, Gent M, Laupacis A, et al. A comparison of aspirin with placebo in patients treated with warfarin after heart-valve replacement. *N Engl J Med*. 1993;329(8):524-529.
83. Thrombosis prevention trial: randomised trial of low-intensity oral anticoagulation with warfarin and low-dose aspirin in the primary prevention of ischaemic heart disease in men at increased risk. The Medical Research Council's General Practice Research Framework. *Lancet*. 1998;351(9098):233-241.
84. Schelleman H, Bilker WB, Brensinger CM, Wan F, Yang YX, Hennessy S. Fibrate/Statin initiation in warfarin users and gastrointestinal bleeding risk. *Am J Med*. 2010;123(2): 151-157.
85. Douketis JD, Melo M, Bell CM, Mamdani MM. Does statin therapy decrease the risk for bleeding in patients who are receiving warfarin? *Am J Med*. 2007;120(4):369. e9-369.e14.
86. Holbrook A, Schulman S, Witt DM, et al. Evidence-based management of anticoagulant therapy: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl 1):e152S-e184S.
87. Wittkowsky AK, Bussey HI, Walker MB, Frei CR. Dietary supplement use among anticoagulation clinic patients. *J Thromb Haemost*. 2007;5(4):875-877.
88. Wittkowsky AK. Drug interactions update: drugs, herbs, and oral anticoagulation. *J Thromb Thrombolysis*. 2001;12(1): 67-71.
89. Greenblatt DJ, von Moltke LL. Interaction of warfarin with drugs, natural substances, and foods. *J Clin Pharmacol*. 2005;45(2):127-132.
90. Samuels N. Herbal remedies and anticoagulant therapy. *Thromb Haemost*. 2005;93(1):3-7.
91. Izzo AA, Di Carlo G, Borrelli F, Ernst E. Cardiovascular pharmacotherapy and herbal medicines: the risk of drug interaction. *Int J Cardiol*. 2005;98(1):1-14.
92. Jiang X, Williams KM, Liauw WS, et al. Effect of ginkgo and ginger on the pharmacokinetics and pharmacodynamics of warfarin in healthy subjects. *Br J Clin Pharmacol*. 2005;59(4):425-432.
93. Engelsen J, Nielsen JD, Winther K. Effect of coenzyme Q10 and ginkgo biloba on warfarin dosage in stable, long-term warfarin treated outpatients: a randomized, double blind, placebo-crossover trial. *Thromb Haemost*. 2002;87(6):1075-1076.
94. Yuan CS, Wei G, Dey L, et al. Brief communication: American ginseng reduces warfarin's effect in healthy patients: a randomized, controlled Trial. *Ann Intern Med*. 2004;141(1):23-27.
95. O'Reilly RA, Rytand DA. "Resistance" to warfarin due to unrecognized vitamin K supplementation. *N Engl J Med*. 1980;303(3):160-161.
96. Suttie JW, Mummah-Schendel LL, Shah DV, Lyle BJ, Greger JL. Vitamin K deficiency from dietary vitamin K restriction in humans. *Am J Clin Nutr*. 1988;47(3):475-480.
97. Sadowski JA, Booth SL, Mann KG, et al, eds. *Structure and Mechanism of Activation of Vitamin K Antagonists*. London, England: Arnold; 1996.
98. Bovill EG, Lawson J, Sadowski J, et al. Mechanisms of vitamin K metabolism and vitamin K-dependent hemostasis: Implications for warfarin therapy. In: Ezekowitz MD, ed. *The Heart as a Source of Systemic Embolisation*. New York, NY: Marcel Dekker; 1992.
99. Booth SL, Charnley JM, Sadowski JA, Saltzman E, Bovill EG, Cushman M. Dietary vitamin K1 and stability of oral anticoagulation: proposal of a diet with constant vitamin K1 content. *Thromb Haemost*. 1997;77(3):504-509.
100. Nutescu EA, Shapiro NL, Ibrahim S, West P. Warfarin and its interactions with foods, herbs and other dietary supplements. *Expert Opin Drug Saf*. 2006;5(3):433-451.
101. Harris JE. Interaction of dietary factors with oral anticoagulants: review and applications. *J Am Diet Assoc*. 1995; 95(5):580-584.
102. Wittkowsky AK. Dietary supplements, herbs and oral anticoagulants: the nature of the evidence. *J Thromb Thrombolysis*. 2008;25(1):72-77.
103. Barcellona D, Contu P, Marongiu F. Patient education and oral anticoagulant therapy. *Haematologica*. 2002;87(10): 1081-1086.
104. Mammen EF. Coagulation abnormalities in liver disease. *Hematol Oncol Clin North Am*. 1992;6(6):1247-1257.
105. Deitcher SR. Interpretation of the international normalised ratio in patients with liver disease. *Lancet*. 2002;359(9300): 47-48.
106. Richards RK. Influence of fever upon the action of 3,3methylene bis-(4-hydroxoycoumarin). *Science*. 1943; 97(2518):313-316.
107. Zevin S, Benowitz NL. Drug interactions with tobacco smoking. An update. *Clin Pharmacokinet*. 1999;36(6):425-438.
108. Evans M, Lewis GM. Increase in international normalized ratio after smoking cessation in a patient receiving warfarin. *Pharmacotherapy*. 2005;25(11):1656-1659.
109. Kuykendall JR, Houle MD, Rhodes RS. Possible warfarin failure due to interaction with smokeless tobacco. *Ann Pharmacother*. 2004;38(4):595-597.
110. Self TH, Reaves AB, Oliphant CS, Sands C. Does heart failure exacerbation increase response to warfarin? A critical

review of the literature. *Curr Med Res Opin*. 2006;22(11):2089-2094.
111. Dreisbach AW, Japa S, Gebrekal AB, et al. Cytochrome P4502C9 activity in end-stage renal disease. *Clin Pharmacol Ther*. 2003;73(5):475-477.
112. Gurwitz JH, Avorn J, Ross-Degnan D, Choodnovskiy I, Ansell J. Aging and the anticoagulant response to warfarin therapy. *Ann Intern Med*. 1992;116(11):901-904.
113. Shepherd AM, Hewick DS, Moreland TA, Stevenson IH. Age as a determinant of sensitivity to warfarin. *Br J Clin Pharmacol*. 1977;4(3):315-320.
114. Hodges SJ, Pilkington MJ, Shearer MJ, Bitensky L, Chayen J. Age-related changes in the circulating levels of congeners of vitamin K2, menaquinone-7 and menaquinone-8. *Clin Sci (Lond)*. 1990;78(1):63-66.
115. Garcia D, Regan S, Crowther M, Hughes RA, Hylek EM. Warfarin maintenance dosing patterns in clinical practice: implications for safer anticoagulation in the elderly population. *Chest*. 2005;127(6):2049-2056.
116. Wessler S, Gitel SN. Warfarin. From bedside to bench. *N Engl J Med*. 1984;311(10):645-652.
117. Zivelin A, Rao LV, Rapaport SI. Mechanism of the anticoagulant effect of warfarin as evaluated in rabbits by selective depression of individual procoagulant vitamin K-dependent clotting factors. *J Clin Invest*. 1993;92(5):2131-2140.
118. Patel P, Weitz J, Brooker LA, Paes B, Mitchell L, Andrew M. Decreased thrombin activity of fibrin clots prepared in cord plasma compared with adult plasma. *Pediatr Res*. 1996;39(5):826-830.
119. Weitz JI, Hudoba M, Massel D, Maraganore J, Hirsh J. Clot-bound thrombin is protected from inhibition by heparin-antithrombin III but is susceptible to inactivation by antithrombin III-independent inhibitors. *J Clin Invest*. 1990;86(2):385-391.
120. Quick A. The prothrombin time in haemophilia and in obstructive jaundice. *J Biol Chem*. 1935;109:73-74.
121. Kirkwood TB. Calibration of reference thromboplastins and standardisation of the prothrombin time ratio. *Thromb Haemost*. 1983;49(3):238-244.
122. van den Besselaar AM, Chantarangkul V, Tripodi A. Thromboplastin standards. *Biologicals*. 2010;38(4):430-436.
123. World Health Organization Expert Committee on Biological Standardization. 33rd Report: Technical Report Series No. 687. Geneva, Switzerland: World Health Organization. 1983
124. Johnston M, Harrison L, Moffat K, Willan A, Hirsh J. Reliability of the international normalized ratio for monitoring the induction phase of warfarin: comparison with the prothrombin time ratio. *J Lab Clin Med*. 1996;128(2):214-217.
125. Tripodi A, Chantarangkul V, Primignani M, et al. The international normalized ratio calibrated for cirrhosis (INR(liver)) normalizes prothrombin time results for model for end-stage liver disease calculation. *Hepatology*. 2007;46(2):520-527.
126. Bellest L, Eschwège V, Poupon R, Chazouillères O, Robert A. A modified international normalized ratio as an effective way of prothrombin time standardization in hepatology. *Hepatology*. 2007;46(2):528-534.
127. Lind SE, Pearce LA, Feinberg WM, Bovill EG. Clinically significant differences in the International Normalized Ratio measured with reagents of different sensitivities. SPAF Investigators. Stroke Prevention in Atrial Fibrillation. *Blood Coagul Fibrinolysis*. 1999;10(5):215-227.
128. Poggio M, van den Besselaar AMHP, van der Velde EA, Bertina RM. The effect of some instruments for prothrombin time testing on the International Sensitivity Index (ISI) of two rabbit tissue thromboplastin reagents. *Thromb Haemost*. 1989;62(3):868-874.
129. D'Angelo A, Seveso MP, D'Angelo SV, et al. Comparison of two automated coagulometers and the manual tilt-tube method for the determination of prothrombin time. *Am J Clin Pathol*. 1989;92(3):321-328.
130. Poller L, Thomson JM, Taberner DA. Effect of automation on prothrombin time test in NEQAS surveys. *J Clin Pathol*. 1989;42(1):97-100.
131. Ray MJ, Smith IR. The dependence of the International Sensitivity Index on the coagulometer used to perform the prothrombin time. *Thromb Haemost*. 1990;63(3):424-429.
132. van Rijn JLML, Schmidt NA, Rutten W. Correction of instrument and reagent based differences in determination of the international normalised ratio (INR) for monitoring anticoagulant therapy. *Clin Chem*. 1989;35(5):840-843.
133. Thomson JM, Taberner DA, Poller L. Automation and prothrombin time: a United Kingdom field study of two widely used coagulometers. *J Clin Pathol*. 1990;43(8):679-684.
134. Finck KM, Doetkott C, Miller DR. Clinical impact of interlaboratory variation in international normalized ratio determinations. *Am J Health Syst Pharm*. 2001;58(8):684-688.
135. Jacobson AK, Ruybalid RL, Johnston M, et al. Significant variation in the reporting of prothrombin time results despite utilization of the INR method of reporting [abstract]. *Circulation*. 1999;100:I-620.
136. Fairweather RB, Ansell J, van den Besselaar AM, et al. College of American Pathologists Conference XXXI on laboratory monitoring of anticoagulant therapy: laboratory monitoring of oral anticoagulant therapy. *Arch Pathol Lab Med*. 1998;122(9):768-781.
137. Ng VL, Levin J, Corash L, Gottfried EL. Failure of the International Normalized Ratio to generate consistent results within a local medical community. *Am J Clin Pathol*. 1993;99(6):689-694.
138. Poller L. Laboratory control of oral anticoagulants. *Br Med J (Clin Res Ed)*. 1987;294(6581):1184.
139. Kazama M, Suzuki S, Abe T, et al. Evaluation of international normalized ratios by a controlled field survey with 4 different thromboplastin reagents. *Thromb Haemost*. 1990;64(4):535-541.
140. van den Besselaar AM, Lewis SM, Mannucci PM, Poller L. Status of present and candidate International Reference Preparations (IRP) of thromboplastin for the prothrombin time: a report of the Subcommittee for Control of Anticoagulation. *Thromb Haemost*. 1993;69(1):85.
141. Duncan EM, Casey CR, Duncan BM, Lloyd JV. Effect of concentration of trisodium citrate anticoagulant on calculation of the International Normalised Ratio and the International Sensitivity Index of thromboplastin. *Thromb Haemost*. 1994;72(1):84-88.
142. Adcock DM, Kressin DC, Marlar RA. Effect of 3.2% vs 3.8% sodium citrate concentration on routine coagulation testing. *Am J Clin Pathol*. 1997;107(1):105-110.
143. Harrison L, Johnston M, Massicotte MP, Crowther M, Moffat K, Hirsh J. Comparison of 5-mg and 10-mg loading doses in initiation of warfarin therapy. *Ann Intern Med*. 1997;126(2):133-136.
144. O'Reilly RA, Aggeler PM. Studies on coumarin anticoagulant drugs. Initiation of warfarin therapy without a loading dose. *Circulation*. 1968;38(1):169-177.
145. Crowther MA, Ginsberg JB, Kearon C, et al. A randomized trial comparing 5-mg and 10-mg warfarin loading doses. *Arch Intern Med*. 1999;159(1):46-48.
146. Kovacs MJ, Rodger M, Anderson DR, et al. Comparison of 10-mg and 5-mg warfarin initiation nomograms together with low-molecular-weight heparin for outpatient treatment of acute venous thromboembolism. A randomized,

double-blind, controlled trial. *Ann Intern Med*. 2003;138(9):714-719.
147. Siguret V, Gouin I, Debray M, et al. Initiation of warfarin therapy in elderly medical inpatients: a safe and accurate regimen. *Am J Med*. 2005;118(2):137-142.
148. O'Connell MB, Kowal PR, Allivato CJ, Repka TL. Evaluation of warfarin initiation regimens in elderly inpatients. *Pharmacotherapy*. 2000;20(8):923-930.
149. Ageno W, Turpie AGG, Steidl L, et al. Comparison of a daily fixed 2.5-mg warfarin dose with a 5-mg, international normalized ratio adjusted, warfarin dose initially following heart valve replacement. *Am J Cardiol*. 2001;88(1):40-44.
150. Vecsler M, Loebstein R, Almog S, et al. Combined genetic profiles of components and regulators of the vitamin K-dependent gamma-carboxylation system affect individual sensitivity to warfarin. *Thromb Haemost*. 2006;95(2):205-211.
151. Herman D, Peternel P, Stegnar M, Breskvar K, Dolzan V. The influence of sequence variations in factor VII, gamma-glutamyl carboxylase and vitamin K epoxide reductase complex genes on warfarin dose requirement. *Thromb Haemost*. 2006;95(5):782-787.
152. Limdi NA, Arnett DK, Goldstein JA, et al. Influence of CYP2C9 and VKORC1on warfarin dose, anticoagulation attainment and maintenance among European-Americans and African-Americans. *Pharmacogenomics*. 2008;9(5):511-526.
153. Voora D, Eby C, Linder MW, et al. Prospective dosing of warfarin based on cytochrome P-450 2C9 genotype. *Thromb Haemost*. 2005;93(4):700-705.
154. Gage BF, Eby C, Johnson JA, et al. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. *Clin Pharmacol Ther*. 2008;84(3):326-331.
155. Anderson JL, Horne BD, Stevens SM, et al; Couma-Gen Investigators. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. *Circulation*. 2007;116(22):2563-2570.
156. Caraco Y, Blotnick S, Muszkat M. CYP2C9 genotype-guided warfarin prescribing enhances the efficacy and safety of anticoagulation: a prospective randomized controlled study. *Clin Pharmacol Ther*. 2008;83(3):460-470.
157. Hillman MA, Wilke RA, Yale SH, et al. A prospective, randomized pilot trial of model-based warfarin dose initiation using CYP2C9 genotype and clinical data. *Clin Med Res*. 2005;3(3):137-145.
158. Klein TE, Altman RB, Eriksson N, et al; International Warfarin Pharmacogenetics Consortium. Estimation of the warfarin dose with clinical and pharmacogenetic data [published correction appears in *N Engl J Med* 2009;361(16):1613]. *N Engl J Med*. 2009;360(8):753-764.
159. Joffe HV, Xu R, Johnson FB, Longtine J, Kucher N, Goldhaber SZ. Warfarin dosing and cytochrome P450 2C9 polymorphisms. *Thromb Haemost*. 2004;91(6):1123-1128.
160. Limdi NA, Wiener H, Goldstein JA, Acton RT, Beasley TM. Influence of CYP2C9 and VKORC1 on warfarin response during initiation of therapy. *Blood Cells Mol Dis*. 2009;43(1):119-128.
161. Palareti G, Legnani C, Guazzaloca G, et al; ad hoc Study Group of the Italian Federation of Anticoagulation Clinics*. Risks factors for highly unstable response to oral anticoagulation: a case-control study. *Br J Haematol*. 2005;129(1):72-78.
162. Forfar JC. Prediction of hemorrhage during long-term oral coumarin anticoagulation by excessive prothrombin ratio. *Am Heart J*. 1982;103(3):445-446.
163. White HD, Gruber M, Feyzi J, et al. Comparison of outcomes among patients randomized to warfarin therapy according to anticoagulant control: results from SPORTIF III and V. *Arch Intern Med*. 2007;167(3):239-245.
164. Cannegieter SC, Rosendaal FR, Wintzen AR, van der Meer FJ, Vandenbroucke JP, Briët E. Optimal oral anticoagulant therapy in patients with mechanical heart valves. *N Engl J Med*. 1995;333(1):11-17.
165. Connolly SJ, Pogue J, Eikelboom J, et al. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation*. 2008;118(20):2029-2037.
166. Petersen P, Boysen G, Godtfredsen J, Andersen ED, Andersen B. Placebo-controlled, randomised trial of warfarin and aspirin for prevention of thromboembolic complications in chronic atrial fibrillation. The Copenhagen AFASAK study. *Lancet*. 1989;1(8631):175-179.
167. The Boston Area Anticoagulation Trial for Atrial Fibrillation Investigators. The effect of low-dose warfarin on the risk of stroke in patients with nonrheumatic atrial fibrillation. *N Engl J Med*. 1990;323(22):1505-1511.
168. Ezekowitz MD, Bridgers SL, James KE, et al; Veterans Affairs Stroke Prevention in Nonrheumatic Atrial Fibrillation Investigators. Warfarin in the prevention of stroke associated with nonrheumatic atrial fibrillation. *N Engl J Med*. 1992;327(20):1406-1412.
169. The European Atrial Fibrillation Trial Study Group. Optimal oral anticoagulant therapy in patients with nonrheumatic atrial fibrillation and recent cerebral ischemia. *N Engl J Med*. 1995;333(1):5-10.
170. van der Meer FJM, Rosendaal FR, Vandenbroucke JP, Briët E. Bleeding complications in oral anticoagulant therapy. An analysis of risk factors. *Arch Intern Med*. 1993;153(13):1557-1562.
171. Charney R, Leddomado E, Rose DN, Fuster V. Anticoagulation clinics and the monitoring of anticoagulant therapy. *Int J Cardiol*. 1988;18(2):197-206.
172. Wilson DB, Dunn MI, Hassanein K. Low-intensity anticoagulation in mechanical cardiac prosthetic valves. *Chest*. 1991;100(6):1553-1557.
173. Connolly SJ, Laupacis A, Gent M, Roberts RS, Cairns JA, Joyner C. Canadian atrial fibrillation anticoagulation (CAFA) study. *J Am Coll Cardiol*. 1991;18(2):349-355.
174. Hylek EM, Skates SJ, Sheehan MA, Singer DE. An analysis of the lowest effective intensity of prophylactic anticoagulation for patients with nonrheumatic atrial fibrillation. *N Engl J Med*. 1996;335(8):540-546.
175. Palareti G, Legnani C, Cosmi B, Guazzaloca G, Cini M, Mattarozzi S. Poor anticoagulation quality in the first 3 months after unprovoked venous thromboembolism is a risk factor for long-term recurrence. *J Thromb Haemost*. 2005;3(5):955-961.
176. Schmitt L, Speckman J, Ansell J. Quality assessment of anticoagulation dose management: comparative evaluation of measures of time-in-therapeutic range. *J Thromb Thrombolysis*. 2003;15(3):213-216.
177. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thromb Haemost*. 1993;69(3):236-239.
178. Meier DJ, Seva S, Fay WP. A comparison of anticoagulation results of patients managed with narrow vs. standard international normalized ratio target ranges. *J Thromb Haemost*. 2007;5(6):1332-1334.
179. Gadisseur AP, van der Meer FJM, Adriaansen HJ, Fihn SD, Rosendaal FR. Therapeutic quality control of oral anticoagulant therapy comparing the short-acting acenocoumarol and the long-acting phenprocoumon. *Br J Haematol*. 2002;117(4):940-946.
180. Pattacini C, Manotti C, Pini M, Quintavalla R, Dettori AG. A comparative study on the quality of oral anticoagulant

therapy (warfarin versus acenocoumarol). *Thromb Haemost*. 1994;71(2):188-191.
181. Redwood M, Taylor C, Bain BJ, Matthews JH. The association of age with dosage requirement for warfarin. *Age Ageing*. 1991;20(3):217-220.
182. James AH, Britt RP, Raskino CL, Thompson SG. Factors affecting the maintenance dose of warfarin. *J Clin Pathol*. 1992;45(8):704-706.
183. Mungall D, White R. Aging and warfarin therapy. *Ann Intern Med*. 1992;117(10):878-879.
184. Pengo V, Barbero F, Biasiolo A, Pegoraro C, Cucchini U, Iliceto S. A comparison between six- and four-week intervals in surveillance of oral anticoagulant treatment. *Am J Clin Pathol*. 2003;120(6):944-947.
185. Shalev V, Rogowski O, Shimron O, et al. The interval between prothrombin time tests and the quality of oral anticoagulants treatment in patients with chronic atrial fibrillation. *Thromb Res*. 2007;120(2):201-206.
186. Witt DM, Delate T, Clark NP, et al; Warfarin Associated Research Projects and other EnDeavors (WARPED) Consortium. Outcomes and predictors of very stable INR control during chronic anticoagulation therapy. *Blood*. 2009; 114(5):952-956.
187. Witt DM, Delate T, Clark NP, et al; Warped Consortium. Twelve-month outcomes and predictors of very stable INR control in prevalent warfarin users. *J Thromb Haemost*. 2010; 8(4):744-749.
188. Shibata Y, Hashimoto H, Kurata C, Ohno R, Kazui T, Takinami M. Influence of physical activity on warfarin therapy. *Thromb Haemost*. 1998;80(1):203-204.
189. Lenz TL, Lenz NJ, Faulkner MA. Potential interactions between exercise and drug therapy. *Sports Med*. 2004;34(5): 293-306.
190. Sorano GG, Biondi G, Conti M, Mameli G, Licheri D, Marongiu F. Controlled vitamin K content diet for improving the management of poorly controlled anticoagulated patients: a clinical practice proposal. *Haemostasis*. 1993; 23(2):77-82.
191. Sconce E, Khan T, Mason J, Noble F, Wynne H, Kamali F. Patients with unstable control have a poorer dietary intake of vitamin K compared to patients with stable control of anticoagulation. *Thromb Haemost*. 2005;93(5):872-875.
192. Kurnik D, Loebstein R, Rabinovitz H, Austerweil N, Halkin H, Almog S. Over-the-counter vitamin K1-containing multivitamin supplements disrupt warfarin anticoagulation in vitamin K1-depleted patients. A prospective, controlled trial. *Thromb Haemost*. 2004;92(5):1018-1024.
193. Schurgers LJ, Shearer MJ, Hamulyák K, Stöcklin E, Vermeer C. Effect of vitamin K intake on the stability of oral anticoagulant treatment: dose-response relationships in healthy subjects. *Blood*. 2004;104(9):2682-2689.
194. Reese AM, Farnett LE, Lyons RM, Patel B, Morgan L, Bussey HI. Low-dose vitamin K to augment anticoagulation control. *Pharmacotherapy*. 2005;25(12):1746-1751.
195. Ford SK, Misita CP, Shilliday BB, Malone RM, Moore CG, Moll S. Prospective study of supplemental vitamin K therapy in patients on oral anticoagulants with unstable international normalized ratios. *J Thromb Thrombolysis*. 2007;24(1): 23-27.
196. Rombouts EK, Rosendaal FR, van der Meer FJ. Influence of dietary vitamin K intake on subtherapeutic oral anticoagulant therapy. *Br J Haematol*. 2010;149(4):598-605.
197. Sconce E, Avery P, Wynne H, Kamali F. Vitamin K supplementation can improve stability of anticoagulation for patients with unexplained variability in response to warfarin. *Blood*. 2007;109(6):2419-2423.
198. Rombouts EK, Rosendaal FR, Van Der Meer FJ. Daily vitamin K supplementation improves anticoagulant stability. *J Thromb Haemost*. 2007;5(10):2043-2048.
199. de Assis MC, Rabelo ER, Avila CW, Polanczyk CA, Rohde LE. Improved oral anticoagulation after a dietary vitamin k-guided strategy: a randomized controlled trial. *Circulation*. 2009;120(12):1115-1122.
200. Kimmel SE, Chen Z, Price M, et al. The influence of patient adherence on anticoagulation control with warfarin: results from the International Normalized Ratio Adherence and Genetics (IN-RANGE) Study. *Arch Intern Med*. 2007; 167(3):229-235.
201. Orensky IA, Holdford DA. Predictors of noncompliance with warfarin therapy in an outpatient anticoagulation clinic. *Pharmacotherapy*. 2005;25(12):1801-1808.
202. Platt AB, Localio AR, Brensinger CM, et al. Risk factors for nonadherence to warfarin: results from the IN-RANGE study. *Pharmacoepidemiol Drug Saf*. 2008;17(9):853-860.
203. Beyth RJ, Quinn LM, Landefeld CS. Prospective evaluation of an index for predicting the risk of major bleeding in outpatients treated with warfarin. *Am J Med*. 1998;105(2):91-99.
204. Gitter MJ, Jaeger TM, Petterson TM, Gersh BJ, Silverstein MD. Bleeding and thromboembolism during anticoagulant therapy: a population-based study in Rochester, Minnesota. *Mayo Clin Proc*. 1995;70(8):725-733.
205. Steffensen FH, Kristensen K, Ejlersen E, Dahlerup JF, Sørensen HT. Major haemorrhagic complications during oral anticoagulant therapy in a Danish population-based cohort. *J Intern Med*. 1997;242(6):497-503.
206. Willey VJ, Bullano MF, Hauch O, et al. Management patterns and outcomes of patients with venous thromboembolism in the usual community practice setting. *Clin Ther*. 2004;26(7):1149-1159.
207. Palareti G, Leali N, Coccheri S, et al; Italian Study on Complications of Oral Anticoagulant Therapy. Bleeding complications of oral anticoagulant treatment: an inception-cohort, prospective collaborative study (ISCOAT). *Lancet*. 1996;348(9025):423-428.
208. Palareti G, Manotti C, DAngelo A, et al. Thrombotic events during oral anticoagulant treatment: results of the inception-cohort, prospective, collaborative ISCOAT study: ISCOAT study group (Italian Study on Complications of Oral Anticoagulant Therapy). *Thromb Haemost*. 1997;78(6):1438-1443.
209. Abdelhafiz AH, Wheeldon NM. Results of an open-label, prospective study of anticoagulant therapy for atrial fibrillation in an outpatient anticoagulation clinic. *Clin Ther*. 2004;26(9):1470-1478.
210. Veeger NJGM, Piersma-Wichers M, Tijssen JGP, Hillege HL, van der Meer J. Individual time within target range in patients treated with vitamin K antagonists: main determinant of quality of anticoagulation and predictor of clinical outcome. A retrospective study of 2300 consecutive patients with venous thromboembolism. *Br J Haematol*. 2005;128(4):513-519.
211. Cortelazzo S, Finazzi G, Viero P, et al. Thrombotic and hemorrhagic complications in patients with mechanical heart valve prosthesis attending an anticoagulation clinic. *Thromb Haemost*. 1993;69(4):316-320.
212. Chiquette E, Amato MG, Bussey HI. Comparison of an anticoagulation clinic with usual medical care: anticoagulation control, patient outcomes, and health care costs. *Arch Intern Med*. 1998;158(15):1641-1647.
213. Witt DM, Sadler MA, Shanahan RL, Mazzoli G, Tillman DJ. Effect of a centralized clinical pharmacy anticoagulation service on the outcomes of anticoagulation therapy. *Chest*. 2005;127(5):1515-1522.

214. Matchar DB, Samsa GP, Cohen SJ, Oddone EZ, Jurgelski AE. Improving the quality of anticoagulation of patients with atrial fibrillation in managed care organizations: results of the managing anticoagulation services trial. *Am J Med*. 2002;113(1):42-51.
215. Wilson SJ, Wells PS, Kovacs MJ, et al. Comparing the quality of oral anticoagulant management by anticoagulation clinics and by family physicians: a randomized controlled trial. *CMAJ*. 2003;169(4):293-298.
216. Ansell J, Hollowell J, Pengo V, Martinez-Brotons F, Caro J, Drouet L. Descriptive analysis of the process and quality of oral anticoagulation management in real-life practice in patients with chronic non-valvular atrial fibrillation: the international study of anticoagulation management (ISAM). *J Thromb Thrombolysis*. 2007;23(2):83-91.
217. van Walraven C, Jennings A, Oake N, Fergusson D, Forster AJ. Effect of study setting on anticoagulation control: a systematic review and metaregression. *Chest*. 2006;129(5):1155-1166.
218. Wilson R, James AH. Computer assisted management of warfarin treatment. *Br Med J (Clin Res Ed)*. 1984;289(6442):422-424.
219. Ryan PJ, Gilbert M, Rose PE. Computer control of anticoagulant dose for therapeutic management. *BMJ*. 1989;299(6709):1207-1209.
220. Weston Smith SG, Savidge GF. Computer control of anticoagulant dose. *BMJ*. 1989;299:1529.
221. Poller L, Wright D, Rowlands M. Prospective comparative study of computer programs used for management of warfarin. *J Clin Pathol*. 1993;46(4):299-303.
222. Ageno W, Turpie AGG, Steidl L, et al. A randomized comparison of a computer-based dosing program with a manual system to monitor oral anticoagulant therapy. *Thromb Res*. 1998;91(5):237-240.
223. Poller L, Shiach CR, MacCallum P, et al. Multicentre randomized study of computerized anticoagulant dosage. European Concerted Action on Anticoagulation. *Lancet*. 1998;352(9139):1505-1509.
224. Manotti C, Moia M, Palareti G, Pengo V, Ria L, Dettori AG. Effect of computer-aided management on the quality of treatment in anticoagulated patients: a prospective, randomized, multicenter trial of APROAT (Automated PRogram for Oral Anticoagulant Treatment). *Haematologica*. 2001;86(10):1060-1070.
225. Poller L, Keown M, Ibrahim S, et al. An international multicenter randomized study of computer-assisted oral anticoagulant dosage vs. medical staff dosage. *J Thromb Haemost*. 2008;6(6):935-943.
226. Jowett S, Bryan S, Poller L, et al. The cost-effectiveness of computer-assisted anticoagulant dosage: results from the European Action on Anticoagulation (EAA) multicentre study. *J Thromb Haemost*. 2009;7(9):1482-1490.
227. Ageno W, Johnson J, Nowacki B, Turpie AG. A computer generated induction system for hospitalized patients starting on oral anticoagulant therapy. *Thromb Haemost*. 2000;83(6):849-852.
228. Ansell JE, Leaning KE, eds. *Capillary Whole Blood Prothrombin Time Monitoring: Instrumentation and Methodologies*. New York, NY: Aspen; 1997.
229. Lucas FV, Duncan A, Jay R, et al. A novel whole blood capillary technic for measuring the prothrombin time. *Am J Clin Pathol*. 1987;88(4):442-446.
230. Yano Y, Kambayashi J, Murata K, et al. Bedside monitoring of warfarin therapy by a whole blood capillary coagulation monitor. *Thromb Res*. 1992;66(5):583-590.
231. Weibert RT, Adler DS. Evaluation of a capillary whole-blood prothrombin time measurement system. *Clin Pharm*. 1989;8(12):864-867.
232. Rose VL, Dermott SC, Murray BF, McIver MM, High KA, Oberhardt BJ. Decentralized testing for prothrombin time and activated partial thromboplastin time using a dry chemistry portable analyzer. *Arch Pathol Lab Med*. 1993;117(6):611-617.
233. Gosselin R, Owings JT, White RH, et al. A comparison of point-of-care instruments designed for monitoring oral anticoagulation with standard laboratory methods. *Thromb Haemost*. 2000;83(5):698-703.
234. van den Besselaar AM. A comparison of INRs determined with a whole blood prothrombin time device and two international reference preparations for thromboplastin. *Thromb Haemost*. 2000;84(3):410-412.
235. Kitchen S, Preston FE. Monitoring oral anticoagulant treatment with the TAS near-patient test system: comparison with conventional thromboplastins. *J Clin Pathol*. 1997;50(11):951-956.
236. Douketis JD, Lane A, Milne J, Ginsberg JS. Accuracy of a portable International Normalization Ratio monitor in outpatients receiving long-term oral anticoagulant therapy: comparison with a laboratory reference standard using clinically relevant criteria for agreement. *Thromb Res*. 1998;92(1):11-17.
237. Cosmi B, Palareti G, Moia M, et al. Accuracy of a portable prothrombin time monitor (Coagucheck) in patients on chronic oral anticoagulant therapy: a prospective multicenter study. *Thromb Res*. 2000;100(4):279-286.
238. Oral Anticoagulation Monitoring Study Group. Prothrombin measurement using a patient self-testing system. *Am J Clin Pathol*. 2001;115(2):280-287.
239. Oral Anticoagulation Monitoring Study Group. Point-of-care prothrombin time measurement for professional and patient self-testing use. A multicenter clinical experience. *Am J Clin Pathol*. 2001;115(2):288-296.
240. Cachia PG, McGregor E, Adlakha S, Davey P, Goudie BM. Accuracy and precision of the TAS analyser for near-patient INR testing by non-pathology staff in the community. *J Clin Pathol*. 1998;51(1):68-72.
241. Murray ET, Greaves M. INRs and point of care testing. *BMJ*. 2003;327(7405):5-6.
242. Jennings I, Luddington RJ, Baglin T. Evaluation of the Ciba Corning Biotrack 512 coagulation monitor for the control of oral anticoagulation. *J Clin Pathol*. 1991;44(11):950-953.
243. McCurdy SA, White RH. Accuracy and precision of a portable anticoagulation monitor in a clinical setting. *Arch Intern Med*. 1992;152(3):589-592.
244. Tripodi A, Arbini AA, Chantarangkul V, Bettega D, Mannucci PM. Are capillary whole blood coagulation monitors suitable for the control of oral anticoagulant treatment by the international normalized ratio? *Thromb Haemost*. 1993;70(6):921-924.
245. Plesch W, Wolf T, Breitenbeck N, et al. Results of the performance verification of the CoaguChek XS system. *Thromb Res*. 2008;123(2):381-389.
246. Ryan F, O'Shea S, Byrne S. The reliability of point-of-care prothrombin time testing. A comparison of CoaguChek S and XS INR measurements with hospital laboratory monitoring. *Int J Lab Hematol*. 2010;32(1 pt 1):e26-e33.
247. Plesch W, van den Besselaar AM. Validation of the international normalized ratio (INR) in a new point-of-care system designed for home monitoring of oral anticoagulation therapy. *Int J Lab Hematol*. 2009;31(1):20-25.
248. Yelland LN, Gialamas A, Laurence CO, Willson KJ, Ryan P, Beilby JJ; PoCT Trial Management Committee. Assessing agreement between point of care and pathology laboratory results for INR: experiences from the Point of Care

Testing in General Practice Trial. *Pathology*. 2010;42(2):155-159.
249. Tripodi A, Chantarangkul V, Clerici M, Negri B, Mannucci PM. Determination of the International Sensitivity Index of a new near-patient testing device to monitor oral anticoagulant therapy—overview of the assessment of conformity to the calibration model. *Thromb Haemost*. 1997;78(2):855-858.
250. Perry SL, Samsa GP, Ortel TL. Point-of-care testing of the international normalized ratio in patients with antiphospholipid antibodies. *Thromb Haemost*. 2005;94(6):1196-1202.
251. Gialamas A, St John A, Laurence CO, Bubner TK; PoCT Management Committee. Point-of-care testing for patients with diabetes, hyperlipidaemia or coagulation disorders in the general practice setting: a systematic review. *Fam Pract*. 2010;27(1):17-24.
252. Kaatz SS, White RH, Hill J, Mascha E, Humphries JE, Becker DM. Accuracy of laboratory and portable monitor international normalized ratio determinations. Comparison with a criterion standard. *Arch Intern Med*. 1995;155(17):1861-1867.
253. Poller L, Keown M, Chauhan N, et al. European Concerted Action on Anticoagulation (ECAA): multicentre international sensitivity index calibration of two types of point-of-care prothrombin time monitor systems. *Br J Haematol*. 2002;116(4):844-850.
254. Poller L, Keown M, Chauhan N, et al. European Concerted Action on Anticoagulation (ECAA). An assessment of lyophilised plasmas for ISI calibration of CoaguChek and TAS whole blood prothrombin time monitors. *J Clin Pathol*. 2003;56(2):114-119.
255. Poller L, Keown M, Chauhan N, et al. Reliability of international normalised ratios from two point of care test systems: comparison with conventional methods. *BMJ*. 2003;327(7405):30-34.
256. Kitchen S, Kitchen DP, Jennings I, Woods TAL, Walker ID, Preston FE. Point-of-care International Normalised Ratios: UK NEQAS experience demonstrates necessity for proficiency testing of three different monitors. *Thromb Haemost*. 2006;96(5):590-596.
257. Jespersen J, Poller L, van den Besselaar AM, et al. External quality assessment (EQA) for CoaguChek monitors. *Thromb Haemost*. 2010;103(5):936-941.
258. Ansell JE, Hughes R. Evolving models of warfarin management: anticoagulation clinics, patient self-monitoring, and patient self-management. *Am Heart J*. 1996;132(5):1095-1100.
259. Christensen TD. Self-management of oral anticoagulant therapy: a review. *J Thromb Thrombolysis*. 2004;18(2):127-143.
260. Siebenhofer A, Berghold A, Sawicki PT. Systematic review of studies of self-management of oral anticoagulation. *Thromb Haemost*. 2004;91(2):225-232.
261. Heneghan C, Alonso-Coello P, Garcia-Alamino JM, Perera R, Meats E, Glasziou P. Self-monitoring of oral anticoagulation: a systematic review and meta-analysis. *Lancet*. 2006;367(9508):404-411.
262. Dauphin C, Legault B, Jaffeux P, et al. Comparison of INR stability between self-monitoring and standard laboratory method: preliminary results of a prospective study in 67 mechanical heart valve patients. *Arch Cardiovasc Dis*. 2008;101(11-12):753-761.
263. Eitz T, Schenk S, Fritzsche D, et al. International normalized ratio self-management lowers the risk of thromboembolic events after prosthetic heart valve replacement. *Ann Thorac Surg*. 2008;85(3):949-954.
264. Ansell J, Jacobson A, Levy J, Völler H, Hasenkam JM; International Self-Monitoring Association for Oral Anticoagulation. Guidelines for implementation of patient self-testing and patient self-management of oral anticoagulation. International consensus guidelines prepared by International Self-Monitoring Association for Oral Anticoagulation. *Int J Cardiol*. 2005;99(1):37-45.
265. Fitzmaurice DA, Gardiner C, Kitchen S, Mackie I, Murray ET, Machin SJ; British Society of Haematology Taskforce for Haemostasis and Thrombosis. An evidence-based review and guidelines for patient self-testing and management of oral anticoagulation. *Br J Haematol*. 2005;131(2):156-165.
266. Landefeld CS, Rosenblatt MW, Goldman L. Bleeding in outpatients treated with warfarin: relation to the prothrombin time and important remediable lesions. *Am J Med*. 1989;87(2):153-159.
267. Fihn SD, McDonell M, Martin D, et al. Risk factors for complications of chronic anticoagulation. A multi-center study. Warfarin Optimized Outpatient Follow-up Study Group. *Ann Intern Med*. 1993;118(7):511-520.
268. Hylek EM, Singer DE. Risk factors for intracranial hemorrhage in outpatients taking warfarin. *Ann Intern Med*. 1994;120(11):897-902.
269. Hull R, Hirsh J, Jay R, et al. Different intensities of oral anticoagulant therapy in the treatment of proximal-vein thrombosis. *N Engl J Med*. 1982;307(27):1676-1681.
270. Ridker PM, Goldhaber SZ, Danielson E, et al; PREVENT Investigators. Long-term, low-intensity warfarin therapy for the prevention of recurrent venous thromboembolism. *N Engl J Med*. 2003;348(15):1425-1434.
271. Kearon C, Ginsberg JS, Kovacs MJ, et al; Extended Low-Intensity Anticoagulation for Thrombo-Embolism Investigators. Comparison of low-intensity warfarin therapy with conventional-intensity warfarin therapy for long-term prevention of recurrent venous thromboembolism. *N Engl J Med*. 2003;349(7):631-639.
272. Turpie AGG, Gunstensen J, Hirsh J, Nelson H, Gent M. Randomised comparison of two intensities of oral anticoagulant therapy after tissue heart valve replacement. *Lancet*. 1988;1(8597):1242-1245.
273. Saour JN, Sieck JO, Mamo LAR, Gallus AS. Trial of different intensities of anticoagulation in patients with prosthetic heart valves. *N Engl J Med*. 1990;322(7):428-432.
274. Altman R, Rouvier J, Gurfinkel E, et al. Comparison of two levels of anticoagulant therapy in patients with substitute heart valves. *J Thorac Cardiovasc Surg*. 1991;101(3):427-431.
275. Hering D, Piper C, Bergemann R, et al. Thromboembolic and bleeding complications following St. Jude Medical valve replacement: results of the German Experience With Low-Intensity Anticoagulation Study. *Chest*. 2005;127(1):53-59.
276. Vink R, Kraaijenhagen RA, Hutten BA, et al. The optimal intensity of vitamin K antagonists in patients with mechanical heart valves: a meta-analysis. *J Am Coll Cardiol*. 2003;42(12):2042-2048.
277. Anand SS, Yusuf S. Oral anticoagulant therapy in patients with coronary artery disease: a meta-analysis. *JAMA*. 1999;282(21):2058-2067.
278. Segal JB, McNamara RL, Miller MR, et al. Prevention of thromboembolism in atrial fibrillation. A meta-analysis of trials of anticoagulants and antiplatelet drugs. *J Gen Intern Med*. 2000;15(1):56-67.
279. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. *Ann Intern Med*. 1999;131(7):492-501.
280. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. *Arch Intern Med*. 1994;154(13):1449-1457.

281. The Stroke Prevention in Reversible Ischemia Trial (SPIRIT) Study Group. A randomized trial of anticoagulants versus aspirin after cerebral ischemia of presumed arterial origin. *Ann Neurol*. 1997;42(6):857-865.
282. Stroke Prevention in Atrial Fibrillation III Investigators. Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation: Stroke Prevention in Atrial Fibrillation III randomised clinical trial. *Lancet*. 1996;348(9028):633-638.
283. Hylek EM, Go AS, Chang Y, et al. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N Engl J Med*. 2003;349(11):1019-1026.
284. Indredavik B, Rohweder G, Lydersen S. Frequency and effect of optimal anticoagulation before onset of ischaemic stroke in patients with known atrial fibrillation. *J Intern Med*. 2005;258(2):133-144.
285. O'Donnell M, Oczkowski W, Fang J, et al; Investigators of the Registry of the Canadian Stroke Network. Preadmission antithrombotic treatment and stroke severity in patients with atrial fibrillation and acute ischaemic stroke: an observational study. *Lancet Neurol*. 2006;5(9):749-754.
286. Crowther MA, Clase CM, Margetts PJ, et al. Low-intensity warfarin is ineffective for the prevention of PTFE graft failure in patients on hemodialysis: a randomized controlled trial. *J Am Soc Nephrol*. 2002;13(9):2331-2337.
287. Gulløv AL, Koefoed BG, Petersen P, et al. Fixed minidose warfarin and aspirin alone and in combination vs adjusted-dose warfarin for stroke prevention in atrial fibrillation: Second Copenhagen Atrial Fibrillation, Aspirin, and Anticoagulation Study. *Arch Intern Med*. 1998;158(14):1513-1521.
288. Coumadin Aspirin Reinfarction Study (CARS) Investigators. Randomised, double-blind trial of fixed low-dose warfarin with aspirin after myocardial infarction. *Lancet*. 1997;350(9075):389-396.
289. Pengo V, Zasso A, Barbero F, et al. Effectiveness of fixed minidose warfarin in the prevention of thromboembolism and vascular death in nonrheumatic atrial fibrillation. *Am J Cardiol*. 1998;82(4):433-437.
290. Couban S, Goodyear M, Burnell M, et al. Randomized placebo-controlled study of low-dose warfarin for the prevention of central venous catheter-associated thrombosis in patients with cancer. *J Clin Oncol*. 2005;23(18):4063-4069.
291. Garcia DA, Regan S, Crowther M, Hylek EM. The risk of hemorrhage among patients with warfarin-associated coagulopathy. *J Am Coll Cardiol*. 2006;47(4):804-808.
292. Clark NP, Witt DM, Delate T, et al; Warfarin-Associated Research Projects and Other Endeavors Consortium. Thromboembolic consequences of subtherapeutic anticoagulation in patients stabilized on warfarin therapy: the low INR study. *Pharmacotherapy*. 2008;28(8):960-967.
293. Dentali F, Riva N, Malato A, Saccullo G, Siragusa S, Ageno W. Incidence of thromboembolic complications in patients with mechanical heart valves with a subtherapeutic international normalized ratio. *J Thorac Cardiovasc Surg*. 2009;137(1):91-93.
294. Garcia DA, Regan S, Henault LE, et al. Risk of thromboembolism with short-term interruption of warfarin therapy. *Arch Intern Med*. 2008;168(1):63-69.
295. Delate T, Witt DM, Jones JR, Bhardwaja B, Senser M. Falsely elevated international normalized ratio values in patients undergoing anticoagulation therapy: a descriptive evaluation. *Chest*. 2007;131(3):816-822.
296. Coon WW, Willis PW III. Hemorrhagic complications of anticoagulant therapy. *Arch Intern Med*. 1974;133(3):386-392.
297. Jaffin BW, Bliss CM, LaMont JT. Significance of occult gastrointestinal bleeding during anticoagulation therapy. *Am J Med*. 1987;83(2):269-272.
298. Wilcox CM, Truss CD. Gastrointestinal bleeding in patients receiving long-term anticoagulant therapy. *Am J Med*. 1988;84(4):683-690.
299. Wolf AT, Wasan SK, Saltzman JR. Impact of anticoagulation on rebleeding following endoscopic therapy for nonvariceal upper gastrointestinal hemorrhage. *Am J Gastroenterol*. 2007;102(2):290-296.
300. Culclasure TF, Bray VJ, Hasbargen JA. The significance of hematuria in the anticoagulated patient. *Arch Intern Med*. 1994;154(6):649-652.
301. Schuster GA, Lewis GA. Clinical significance of hematuria in patients on anticoagulant therapy. *J Urol*. 1987;137(5):923-925.
302. Van Savage JG, Fried FA. Anticoagulant associated hematuria: a prospective study. *J Urol*. 1995;153(5):1594-1596.
303. Garcia D, Ageno W, Bussey H, et al. Prevention and treatment of bleeding complications in patients receiving vitamin K antagonists, Part 1: Prevention. *Am J Hematol*. 2009;84(9):579-583.
304. Dahri K, Loewen P. The risk of bleeding with warfarin: a systematic review and performance analysis of clinical prediction rules. *Thromb Haemost*. 2007;98(5):980-987.
305. Gage BF, Yan Y, Milligan PE, et al. Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). *Am Heart J*. 2006;151(3):713-719.
306. Wickramasinghe LSP, Basu SK, Bansal SK. Long-term oral anticoagulant therapy in elderly patients. *Age Ageing*. 1988;17(6):388-396.
307. Copland M, Walker ID, Tait RC. Oral anticoagulation and hemorrhagic complications in an elderly population with atrial fibrillation. *Arch Intern Med*. 2001;161(17):2125-2128.
308. Launbjerg J, Egeblad H, Heaf J, Nielsen NH, Fugleholm AM, Ladefoged K. Bleeding complications to oral anticoagulant therapy: multivariate analysis of 1010 treatment years in 551 outpatients. *J Intern Med*. 1991;229(4):351-355.
309. The Stroke Prevention in Atrial Fibrillation Investigators. Bleeding during antithrombotic therapy in patients with atrial fibrillation. *Arch Intern Med*. 1996;156(4):409-416.
310. Fihn SD, Callahan CM, Martin DC, McDonell MB, Henikoff JG, White RH; The National Consortium of Anticoagulation Clinics. The risk for and severity of bleeding complications in elderly patients treated with warfarin. *Ann Intern Med*. 1996;124(11):970-979.
311. Petitti DB, Strom BL, Melmon KL. Duration of warfarin anticoagulant therapy and the probabilities of recurrent thromboembolism and hemorrhage. *Am J Med*. 1986;81(2):255-259.
312. McCormick D, Gurwitz JH, Goldberg RJ, Ansell J. Long-term anticoagulation therapy for atrial fibrillation in elderly patients: efficacy, risk, and current patterns of use. *J Thromb Thrombolysis*. 1999;7(2):157-163.
313. Cheng HY. Is age a risk factor for warfarin-related major bleeds in elderly patients with atrial fibrillation? *J Am Geriatr Soc*. 2006;54(7):1155.
314. Palareti G, Hirsh J, Legnani C, et al. Oral anticoagulation treatment in the elderly: a nested, prospective, case-control study. *Arch Intern Med*. 2000;160(4):470-478.
315. Fang MC, Go AS, Hylek EM, et al. Age and the risk of warfarin-associated hemorrhage: the anticoagulation and risk factors in atrial fibrillation study. *J Am Geriatr Soc*. 2006;54(8):1231-1236.
316. Dentali F, Douketis JD, Lim W, Crowther M. Combined aspirin-oral anticoagulant therapy compared with oral

anticoagulant therapy alone among patients at risk for cardiovascular disease: a meta-analysis of randomized trials. *Arch Intern Med*. 2007;167(2):117-124.
317. Flaker GC, Gruber M, Connolly SJ, et al; SPORTIF Investigators. Risks and benefits of combining aspirin with anticoagulant therapy in patients with atrial fibrillation: an exploratory analysis of stroke prevention using an oral thrombin inhibitor in atrial fibrillation (SPORTIF) trials. *Am Heart J*. 2006;152(5):967-973.
318. Rubboli A, Halperin JL, Airaksinen KE, et al. Antithrombotic therapy in patients treated with oral anticoagulation undergoing coronary artery stenting. An expert consensus document with focus on atrial fibrillation. *Ann Med*. 2008;40(6):428-436.
319. Hansen ML, Sørensen R, Clausen MT, et al. Risk of bleeding with single, dual, or triple therapy with warfarin, aspirin, and clopidogrel in patients with atrial fibrillation. *Arch Intern Med*. 2010;170(16):1433-1441.
320. Ageno W, Garcia D, Aguilar MI, et al. Prevention and treatment of bleeding complications in patients receiving vitamin K antagonists, part 2: Treatment. *Am J Hematol*. 2009;84(9):584-588.
321. Patel RJ, Witt DM, Saseen JJ, Tillman DJ, Wilkinson DS. Randomized, placebo-controlled trial of oral phytonadione for excessive anticoagulation. *Pharmacotherapy*. 2000;20(10):1159-1166.
322. Fondevila CG, Grosso SH, Santarelli MT, Pinto MD. Reversal of excessive oral anticoagulation with a low oral dose of vitamin K1 compared with acenocoumarine discontinuation. A prospective, randomized, open study. *Blood Coagul Fibrinolysis*. 2001;12(1):9-16.
323. Ageno W, Crowther M, Steidl L, et al. Low dose oral vitamin K to reverse acenocoumarol-induced coagulopathy: a randomized controlled trial. *Thromb Haemost*. 2002;88(1):48-51.
324. Dentali F, Ageno W, Crowther M. Treatment of coumarin-associated coagulopathy: a systematic review and proposed treatment algorithms. *J Thromb Haemost*. 2006;4(9):1853-1863.
325. Lubetsky A, Yonath H, Olchovsky D, Loebstein R, Halkin H, Ezra D. Comparison of oral vs intravenous phytonadione (vitamin K1) in patients with excessive anticoagulation: a prospective randomized controlled study. *Arch Intern Med*. 2003;163(20):2469-2473.
326. Crowther MA, Douketis JD, Schnurr T, et al. Oral vitamin K lowers the international normalized ratio more rapidly than subcutaneous vitamin K in the treatment of warfarin-associated coagulopathy. A randomized, controlled trial. *Ann Intern Med*. 2002;137(4):251-254.
327. Nee R, Doppenschmidt D, Donovan DJ, Andrews TC. Intravenous versus subcutaneous vitamin K1 in reversing excessive oral anticoagulation. *Am J Cardiol*. 1999;83(2):286-288; A6-A7.
328. Raj G, Kumar R, McKinney WP. Time course of reversal of anticoagulant effect of warfarin by intravenous and subcutaneous phytonadione. *Arch Intern Med*. 1999;159(22):2721-2724.
329. Riegert-Johnson DL, Volcheck GW. The incidence of anaphylaxis following intravenous phytonadione (vitamin K1): a 5-year retrospective review. *Ann Allergy Asthma Immunol*. 2002;89(4):400-406.
330. Contreras M, Ala FA, Greaves M, et al. Guidelines for the use of fresh frozen plasma. British Committee for Standards in Haematology, Working Party of the Blood Transfusion Task Force. *Transfus Med*. 1992;2(1):57-63.
331. Popovsky MA. Transfusion-Related Acute Lung Injury: Incidence, Pathogenesis and the Role of Multicomponent Apheresis in Its Prevention. *Transfus Med Hemother*. 2008;35(2):76-79.
332. Leissinger CA, Blatt PM, Hoots WK, Ewenstein B. Role of prothrombin complex concentrates in reversing warfarin anticoagulation: a review of the literature. *Am J Hematol*. 2008;83(2):137-143.
333. Rosovsky RP, Crowther MA. What is the evidence for the off-label use of recombinant factor VIIa (rFVIIa) in the acute reversal of warfarin? ASH evidence-based review 2008. *Hematology Am Soc Hematol Educ Program*. 2008:36-38.
334. Verhagen H. Local hemorrhage and necrosis of the skin and underlying tissues at starting therapy with dicumarol or dicumacyl. *Acta Med Scand*. 1954;148(6):453-467.
335. Weinberg AC, Lieskovsky G, McGehee WG, Skinner DG. Warfarin necrosis of the skin and subcutaneous tissue of the male external genitalia. *J Urol*. 1983;130(2):352-354.
336. Broekmans AW, Bertina RM, Loeliger EA, Hofmann V, Klingemann HG. Protein C and the development of skin necrosis during anticoagulant therapy. *Thromb Haemost*. 1983;49(3):251.
337. Zauber NP, Stark MW. Successful warfarin anticoagulation despite protein C deficiency and a history of warfarin necrosis. *Ann Intern Med*. 1986;104(5):659-660.
338. Samama M, Horellou MH, Soria J, Conard J, Nicolas G. Successful progressive anticoagulation in a severe protein C deficiency and previous skin necrosis at the initiation of oral anticoagulant treatment. *Thromb Haemost*. 1984;51(1):132-133.
339. Grimaudo V, Gueissaz F, Hauert J, Sarraj A, Kruithof EK, Bachmann F. Necrosis of skin induced by coumarin in a patient deficient in protein S. *BMJ*. 1989;298(6668):233-234.
340. Warkentin TE. Venous limb gangrene during warfarin treatment of cancer-associated deep venous thrombosis. *Ann Intern Med*. 2001;135(8 pt 1):589-593.
341. Warkentin TE, Elavathil LJ, Hayward CPM, Johnston MA, Russett JI, Kelton JG. The pathogenesis of venous limb gangrene associated with heparin-induced thrombocytopenia. *Ann Intern Med*. 1997;127(9):804-812.
342. Srinivasan AF, Rice L, Bartholomew JR, et al. Warfarin-induced skin necrosis and venous limb gangrene in the setting of heparin-induced thrombocytopenia. *Arch Intern Med*. 2004;164(1):66-70.
343. Raj K, Collins B, Rangarajan S. Purple toe syndrome following anticoagulant therapy. *Br J Haematol*. 2001;114(4):740.
344. Talmadge DB, Spyropoulos AC. Purple toes syndrome associated with warfarin therapy in a patient with antiphospholipid syndrome. *Pharmacotherapy*. 2003;23(5):674-677.
345. Hauschka PV, Lian JB, Cole DEC, Gundberg CM. Osteocalcin and matrix Gla protein: vitamin K-dependent proteins in bone. *Physiol Rev*. 1989;69(3):990-1047.
346. Price PA. Role of vitamin K-dependent proteins in bone metabolism. *Ann Rev Nutr*. 1988;8:565-583.
347. Maillard C, Berruyer M, Serre CM, Dechavanne M, Delmas PD. Protein-S, a vitamin K-dependent protein, is a bone matrix component synthesized and secreted by osteoblasts. *Endocrinology*. 1992;130(3):1599-1604.
348. Pan LC, Williamson MK, Price PA. Sequence of the precursor to rat bone gamma-carboxyglutamic acid protein that accumulates in warfarin-treated osteosarcoma cells. *J Biol Chem*. 1985;260(25):13398-13401.
349. Pettifor JM, Benson R. Congenital malformations associated with the administration of oral anticoagulants during pregnancy. *J Pediatr*. 1975;86(3):459-462.
350. Hall JG, Pauli RM, Wilson KM. Maternal and fetal sequelae of anticoagulation during pregnancy. *Am J Med*. 1980;68(1):122-140.

351. Barnes C, Newall F, Ignjatovic V, et al. Reduced bone density in children on long-term warfarin. *Pediatr Res*. 2005; 57(4):578-581.
352. Danziger J. Vitamin K-dependent proteins, warfarin, and vascular calcification. *Clin J Am Soc Nephrol*. 2008;3(5): 1504-1510.
353. Lazo-Langner A, Rodger MA, Wells PS. Lessons from ximelagatran: issues for future studies evaluating new oral direct thrombin inhibitors for venous thromboembolism prophylaxis in orthopedic surgery. *Clin Appl Thromb Hemost*. 2009;15(3):316-326.
354. Diener HC; Executive Steering Committee of the SPORTIFF III and V Investigators. Stroke prevention using the oral direct thrombin inhibitor ximelagatran in patients with nonvalvular atrial fibrillation. Pooled analysis from the SPORTIF III and V studies. *Cerebrovasc Dis*. 2006;21(4):279-293.
355. Boudes PF. The challenges of new drugs benefits and risks analysis: lessons from the ximelagatran FDA Cardiovascular Advisory Committee. *Contemp Clin Trials*. 2006;27(5): 432-440.
356. Nutescu EA, Shapiro NL, Chevalier A. New anticoagulant agents: direct thrombin inhibitors. *Cardiol Clin*. 2008;26(2): 169-187; v-vi.
357. Hauel NH, Nar H, Priepke H, Ries U, Stassen JM, Wienen W. Structure-based design of novel potent nonpeptide thrombin inhibitors. *J Med Chem*. 2002;45(9):1757-1766.
358. Wienen W, Stassen JM, Priepke H, Ries UJ, Hauel N. In-vitro profile and ex-vivo anticoagulant activity of the direct thrombin inhibitor dabigatran and its orally active prodrug, dabigatran etexilate. *Thromb Haemost*. 2007;98(1): 155-162.
359. Blech S, Ebner T, Ludwig-Schwellinger E, Stangier J, Roth W. The metabolism and disposition of the oral direct thrombin inhibitor, dabigatran, in humans. *Drug Metab Dispos*. 2008;36(2):386-399.
360. Lange U, Nowak G, Bucha E. Ecarin chromogenic assay—a new method for quantitative determination of direct thrombin inhibitors like hirudin. *Pathophysiol Haemost Thromb*. 2003;33(4):184-191.
361. Stangier J, Rathgen K, Stähle H, Gansser D, Roth W. The pharmacokinetics, pharmacodynamics and tolerability of dabigatran etexilate, a new oral direct thrombin inhibitor, in healthy male subjects. *Br J Clin Pharmacol*. 2007;64(3): 292-303.
362. Wienen W, Stassen JM, Priepke H, Ries UJ, Hauel N. Effects of the direct thrombin inhibitor dabigatran and its orally active prodrug, dabigatran etexilate, on thrombus formation and bleeding time in rats. *Thromb Haemost*. 2007; 98(2):333-338.
363. Wienen W, Stassen JM, Priepke H, Ries UJ, Hauel N. Antithrombotic and anticoagulant effects of the direct thrombin inhibitor dabigatran, and its oral prodrug, dabigatran etexilate, in a rabbit model of venous thrombosis. *J Thromb Haemost*. 2007;5(6):1237-1242.
364. Stangier J, Stähle H, Rathgen K, Fuhr R. Pharmacokinetics and pharmacodynamics of the direct oral thrombin inhibitor dabigatran in healthy elderly subjects. *Clin Pharmacokinet*. 2008;47(1):47-59.
365. Stangier J, Eriksson BI, Dahl OE, et al. Pharmacokinetic profile of the oral direct thrombin inhibitor dabigatran etexilate in healthy volunteers and patients undergoing total hip replacement. *J Clin Pharmacol*. 2005;45(5):555-563.
366. Trocóniz IF, Tillmann C, Liesenfeld KH, Schäfer HG, Stangier J. Population pharmacokinetic analysis of the new oral thrombin inhibitor dabigatran etexilate (BIBR 1048) in patients undergoing primary elective total hip replacement surgery. *J Clin Pharmacol*. 2007;47(3):371-382.
367. Stangier J. Clinical pharmacokinetics and pharmacodynamics of the oral direct thrombin inhibitor dabigatran etexilate. *Clin Pharmacokinet*. 2008;47(5):285-295.
368. Eriksson BI, Dahl OE, Ahnfelt L, et al. Dose escalating safety study of a new oral direct thrombin inhibitor, dabigatran etexilate, in patients undergoing total hip replacement: BISTRO I. *J Thromb Haemost*. 2004;2(9):1573-1580.
369. van Ryn J, Stangier J, Haertter S, et al. Dabigatran etexilate—a novel, reversible, oral direct thrombin inhibitor: interpretation of coagulation assays and reversal of anticoagulant activity. *Thromb Haemost*. 2010;103(6):1116-1127.
370. Stangier J, Rathgen K, Stähle H, Mazur D. Influence of renal impairment on the pharmacokinetics and pharmacodynamics of oral dabigatran etexilate: an open-label, parallel-group, single-centre study. *Clin Pharmacokinet*. 2010; 49(4):259-268.
371. Stangier J, Stähle H, Rathgen K, Roth W, Shakeri-Nejad K. Pharmacokinetics and pharmacodynamics of dabigatran etexilate, an oral direct thrombin inhibitor, are not affected by moderate hepatic impairment. *J Clin Pharmacol*. 2008; 48(12):1411-1419.
372. Ezekowitz MD, Reilly PA, Nehmiz G, et al. Dabigatran with or without concomitant aspirin compared with warfarin alone in patients with nonvalvular atrial fibrillation (PETRO Study). *Am J Cardiol*. 2007;100(9):1419-1426.
373. EMEA. CHMP assessment report for Pradaxa. http://www.emaeuropaeu/humandocs/PDFs/EPAR/pradaxa/H-829-en6pdf. 2008. Accessed March 2, 2011.
374. Mouly S, Meune C, Bergmann JF. Mini-series: I. Basic science. Uncertainty and inaccuracy of predicting CYP-mediated in vivo drug interactions in the ICU from in vitro models: focus on CYP3A4. *Intensive Care Med*. 2009;35(3):417-429.
375. Stangier J, Rathgen K, Stähle H, Reseski K, Körnicke T, Roth W. Coadministration of dabigatran etexilate and atorvastatin: assessment of potential impact on pharmacokinetics and pharmacodynamics. *Am J Cardiovasc Drugs*. 2009;9(1):59-68.
376. Stangier J, Stähle H, Rathgen K, Reseski K, Körnicke T. Coadministration of the oral direct thrombin inhibitor dabigatran etexilate and diclofenac has little impact on the pharmacokinetics of either drug [abstract]. *J Thromb Haemost*. 2007;5(suppl 2):P-T-677.
377. Marchetti S, Mazzanti R, Beijnen JH, Schellens JHM. Concise review: Clinical relevance of drug drug and herb drug interactions mediated by the ABC transporter ABCB1 (MDR1, P-glycoprotein). *Oncologist*. 2007;12(8):927-941.
378. Stangier J, Stähle H, Rathgen K, Reseski K, Körnicke T. No interaction of the oral direct thrombin inhibitor dabigatran etexilate and digoxin [abstract]. *J Thromb Haemost*. 2007;5(suppl 2):P-W-672.
379. Patrono C, García Rodríguez LA, Landolfi R, Baigent C. Low-dose aspirin for the prevention of atherothrombosis. *N Engl J Med*. 2005;353(22):2373-2383.
380. Holmes DR Jr, Kereiakes DJ, Kleiman NS, Moliterno DJ, Patti G, Grines CL. Combining antiplatelet and anticoagulant therapies. *J Am Coll Cardiol*. 2009;54(2):95-109.
381. Connolly SJ, Ezekowitz MD, Yusuf S, et al; RE-LY Steering Committee and Investigators. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med*. 2009; 361(12):1139-1151.
382. Schulman S, Kearon C, Kakkar AK, et al; RE-COVER Study Group. Dabigatran versus warfarin in the treatment of acute venous thromboembolism. *N Engl J Med*. 2009; 361(24):2342-2352.
383. Eriksson BI, Dahl OE, Rosencher N, et al; RE-NOVATE Study Group. Dabigatran etexilate versus enoxaparin for prevention of venous thromboembolism after total hip

replacement: a randomised, double-blind, non-inferiority trial. *Lancet*. 2007;370(9591):949-956.
384. Eriksson BI, Dahl OE, Rosencher N, et al; RE-MODEL Study Group. Oral dabigatran etexilate vs. subcutaneous enoxaparin for the prevention of venous thromboembolism after total knee replacement: the RE-MODEL randomized trial. *J Thromb Haemost*. 2007;5(11):2178-2185.
385. Ginsberg JS, Davidson BL, Comp PC, et al; RE-MOBILIZE Writing Committee. Oral thrombin inhibitor dabigatran etexilate vs North American enoxaparin regimen for prevention of venous thromboembolism after knee arthroplasty surgery. *J Arthroplasty*. 2009;24(1):1-9.
386. Huo MH, Eriksson B, Dahl OE, et al. Oral dabigatran versus enoxaparin for thromboprophylaxis after primary total hip arthroplasty: the RE-NOVATE II randomised trial. Presented at: The European Hematology Association Congress; June 10-13, 2010; Barcelona, Spain.
387. Singer DE, Albers GW, Dalen JE, et al; American College of Chest Physicians. Antithrombotic therapy in atrial fibrillation: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest*. 2008; 133(suppl 6):546S-592S.
388. Schulman S, Beyth RJ, Kearon C, Levine MN; American College of Chest Physicians. Hemorrhagic complications of anticoagulant and thrombolytic treatment: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest*. 2008;133(suppl6): 257S-298S.
389. Mismetti P, Laporte S. New oral antithrombotics: a need for laboratory monitoring. For. *J Thromb Haemost*. 2010; 8(4):621-626.
390. Bounameaux H, Reber G. New oral antithrombotics: a need for laboratory monitoring. Against. *J Thromb Haemost*. 2010; 8(4):627-630.
391. Keisu M, Andersson TB. Drug-induced liver injury in humans: the case of ximelagatran. *Handb Exp Pharmacol*. 2010;196:407-418.
392. Agnelli G, Eriksson BI, Cohen AT, et al; EXTEND Study Group. Safety assessment of new antithrombotic agents: lessons from the EXTEND study on ximelagatran. *Thromb Res*. 2009;123(3):488-497.
393. Eriksson BI, Borris LC, Friedman RJ, et al; RECORD1 Study Group. Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty. *N Engl J Med*. 2008; 358(26):2765-2775.
394. Kakkar AK, Brenner B, Dahl OE, et al; RECORD2 Investigators. Extended duration rivaroxaban versus short-term enoxaparin for the prevention of venous thromboembolism after total hip arthroplasty: a double-blind, randomised controlled trial. *Lancet*. 2008;372(9632):31-39.
395. Lassen MR, Ageno W, Borris LC, et al; RECORD3 Investigators. Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty. *N Engl J Med*. 2008; 358(26):2776-2786.
396. Turpie AG, Lassen MR, Davidson BL, et al; RECORD4 Investigators. Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty (RECORD4): a randomised trial. *Lancet*. 2009;373(9676):1673-1680.
397. Graff J, von Hentig N, Misselwitz F, et al. Effects of the oral, direct factor xa inhibitor rivaroxaban on platelet-induced thrombin generation and prothrombinase activity. *J Clin Pharmacol*. 2007;47(11):1398-1407.
398. Perzborn E, Roehrig S, Straub A, Kubitza D, Mueck W, Laux V. Rivaroxaban: a new oral factor Xa inhibitor. *Arterioscler Thromb Vasc Biol*. 2010;30(3):376-381.
399. Perzborn E, Strassburger J, Wilmen A, et al. In vitro and in vivo studies of the novel antithrombotic agent BAY 59-7939—an oral, direct Factor Xa inhibitor. *J Thromb Haemost*. 2005;3(3):514-521.
400. Gerotziafas GT, Elalamy I, Depasse F, Perzborn E, Samama MM. In vitro inhibition of thrombin generation, after tissue factor pathway activation, by the oral, direct factor Xa inhibitor rivaroxaban. *J Thromb Haemost*. 2007;5(4):886-888.
401. Lang D, Freudenberger C, Weinz C. In vitro metabolism of rivaroxaban, an oral, direct factor Xa inhibitor, in liver microsomes and hepatocytes of rats, dogs, and humans. *Drug Metab Dispos*. 2009;37(5):1046-1055.
402. Samama MM, Martinoli JL, LeFlem L, et al. Assessment of laboratory assays to measure rivaroxaban—an oral, direct factor Xa inhibitor. *Thromb Haemost*. 2010;103(4): 815-825.
403. Kubitza D, Becka M, Mueck W, Zuehlsdorf M. Safety, tolerability, pharmacodynamics, and pharmacokinetics of rivaroxaban—an oral, direct factor Xa inhibitor—are not affected by aspirin. *J Clin Pharmacol*. 2006;46(9):981-990.
404. Kubitza D, Becka M, Voith B, Zuehlsdorf M, Wensing G. Safety, pharmacodynamics, and pharmacokinetics of single doses of BAY 59-7939, an oral, direct factor Xa inhibitor. *Clin Pharmacol Ther*. 2005;78(4):412-421.
405. Kubitza D, Becka M, Wensing G, Voith B, Zuehlsdorf M. Safety, pharmacodynamics, and pharmacokinetics of BAY 59-7939—an oral, direct Factor Xa inhibitor—after multiple dosing in healthy male subjects. *Eur J Clin Pharmacol*. 2005;61(12):873-880.
406. Mueck W, Becka M, Kubitza D, Voith B, Zuehlsdorf M. Population model of the pharmacokinetics and pharmacodynamics of rivaroxaban—an oral, direct factor xa inhibitor—in healthy subjects. *Int J Clin Pharmacol Ther*. 2007;45(6): 335-344.
407. Kubitza D, Becka M, Roth A, Mueck W. Dose-escalation study of the pharmacokinetics and pharmacodynamics of rivaroxaban in healthy elderly subjects. *Curr Med Res Opin*. 2008;24(10):2757-2765.
408. Kubitza D, Becka M, Zuehlsdorf M, Mueck W. Body weight has limited influence on the safety, tolerability, pharmacokinetics, or pharmacodynamics of rivaroxaban (BAY 59-7939) in healthy subjects. *J Clin Pharmacol*. 2007;47(2): 218-226.
409. Kubitza D, Becka M, Mueck W, Zuehlsdorf M. The effect of extreme age, and gender, on the pharmacology and tolerability of rivaroxaban—an oral, direct Factor Xa inhibitor [abstract]. *Blood*. 2006;108:905.
410. Jiang J, Hu Y, Zhang J, et al. Safety, pharmacokinetics and pharmacodynamics of single doses of rivaroxaban—an oral, direct factor Xa inhibitor—in elderly Chinese subjects. *Thromb Haemost*. 2010;103(1):234-241.
411. Zhao X, Sun P, Zhou Y, et al. Safety, pharmacokinetics and pharmacodynamics of single/multiple doses of the oral, direct Factor Xa inhibitor rivaroxaban in healthy Chinese subjects. *Br J Clin Pharmacol*. 2009;68(1):77-88.
412. Mueck W, Eriksson BI, Bauer KA, et al. Population pharmacokinetics and pharmacodynamics of rivaroxaban—an oral, direct factor Xa inhibitor—in patients undergoing major orthopaedic surgery. *Clin Pharmacokinet*. 2008;47(3): 203-216.
413. Mueck W, Borris LC, Dahl OE, et al. Population pharmacokinetics and pharmacodynamics of once- and twice-daily rivaroxaban for the prevention of venous thromboembolism in patients undergoing total hip replacement. *Thromb Haemost*. 2008;100(3):453-461.
414. Bayer Schering Pharma. Xarelto. Summary of Product Characteristics. May 2009. www.xarelto.com/html/downloads/Xarelto_Summary_of_Product_Characteristics_May2009.pdf. Accessed March 2, 2011.

415. Kubitza D, Mueck W, Becka M. No interaction between rivaroxaban—a novel, oral, direct factor Xa inhibitor—and atorvastatin [abstract]. *Pathophysiol Haemost Thromb*. 2008; 36:A40.
416. Kubitza D, Becka M, Zuehlsdorf M. No interaction between the novel, oral direct Factor Xa inhibitor BAY 59-7939 and digoxin [abstract]. *J Clin Pharmacol*. 2006; 46(6):11.
417. Kubitza D, Becka M, Mueck W, Zuehlsdorf M. Rivaroxaban (BAY 59-7939)—an oral, direct Factor Xa inhibitor—has no clinically relevant interaction with naproxen. *Br J Clin Pharmacol*. 2007;63(4):469-476.
418. Kubitza D, Becka M, Zuehlsdorf M, Mueck W. Effect of food, an antacid, and the H2 antagonist ranitidine on the absorption of BAY 59-7939 (rivaroxaban), an oral, direct factor Xa inhibitor, in healthy subjects. *J Clin Pharmacol*. 2006;46(5):549-558.
419. Tersteegen A, Schmidt S, Burkhardt N. Rivaroxaban—an oral, direct Factor Xa inhibitor— binds rapidly to Factor Xa [abstract]. *J Thromb Haemost*. 2007;5(Suppl 2):651.
420. Depasse F, Busson J, Mnich J, Le Flem L, Gerotziafas GT, Samama MM. Effect of BAY 59-7939—a novel, oral, direct Factor Xa inhibitor—on clot-bound Factor Xa activity in vitro [abstract]. *J Thromb Haemost*. 2005;3(suppl 1):P1104.
421. Perzborn E, Strassburger J, Wilmen A, et al. Biochemical and pharmacologic properties of BAY 59-7939, an oral, direct Factor Xa inhibitor [abstract]. *Pathophysiol Haemost Thromb*. 2004;33(suppl 2):79.
422. Hoppensteadt D, Neville B, Schultz C, et al. Interaction of BAY 59-7939—a novel, oral, direct Factor Xa inhibitor—with antiplatelet agents: monitoring and therapeutic applications [abstract]. *J Thromb Haemost*. 2005;3(suppl 1):1717.
423. Perzborn E, Lange U. Rivaroxaban—an oral, direct Factor Xa inhibitor— inhibits tissue factor-mediated platelet aggregation [abstract]. *J Thromb Haemost*. 2007;5(suppl 2):642.
424. Smith SA, Morrissey JH. Thromboplastin composition affects the sensitivity of prothrombin time clotting tests to direct Factor Xa inhibitors [abstract]. *Blood*. 2007;110(11):928.
425. Bauersachs R, Berkowitz SD, Brenner B, et al; EINSTEIN Investigators. Oral rivaroxaban for symptomatic venous thromboembolism. *N Engl J Med*. 2010;363(26):2499-2510.
426. Patel MR, Mahaffey KW, Garg J, et al; ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med*. 2011;365(10):883-891.
427. National Institutes of Health Clinical Center. An efficacy and safety study for rivaroxaban in patients with acute coronary syndrome. NCT00809965. ClinicTrials.gov. Bethesda, MD: National Institutes of Health; 2008. Updated 2011. www.clinicaltrials.gov.
428. Turpie AG, Lassen MR, Eriksson BI, et al. Rivaroxaban for the prevention of venous thromboembolism after hip or knee arthroplasty. Pooled analysis of four studies. *Thromb Haemost*. 2011;105(3):444-453.
429. Perzborn E, Tinel H. Prothrombin complex concentrate reverses the effects of high-dose rivaroxaban in rats [abstract]. *J Thromb Haemost*. 2009;7(suppl 2):183.
430. Lu G, Luan P, Hollenbach SJ, et al. Reconstructed recombinant factor Xa as an antidote to reverse anticoagulation by factor Xa inhibitors [abstract]. *J Thromb Haemost*. 2009; 7(suppl 2):OC-TH-107.

Table S1—*[Section 1.3.1] Observed Frequency of CYP2C9 and VKORC Variants Among Various Ethnic Groups*[1]

CYP2C9 Genetic Alleles	CYP2C9*1	CYP2C9*2	CYP2C9*3	CYP2C9*4	CYP2C9*5
Point mutation	Arg144/IIc359	Cys144/IIc359	Arg144/Leu351	Arg144/Thr359	Arg144/Glu360
Ethnic group, %					
Whites	79-86	8-19.1	6-10	ND	ND
Indigenous Canadians	91	3	6	ND	ND
African Americans	98.5	1-3.6	0.5-1.5	ND	2.3
Asians	95-98.3	0	1.7-5	0-1.6	0
VKORC genetic haplotype sequence	[1]H CCGATCTCTG H2 CCGAGCTCTG	H7 TCGGTCCGCA H8 TAGGTCCGCA H9 TACGTTCGCG			
Ethnic group, %					
Europeans	37	58			
African	14	49			
Asian	89	10			

CYP2C9*2, *3, *4, and *5 represent genetic polymorphisms of the wild-type enzyme, CYP2C9*1. [1]H and H2 represent warfarin-sensitive haplotype. H7, H8, and H9 represent warfarin-resistant haplotype. ND = not determined. (Table adapted from Ansell et al.[1]) (CYP2C9 and VKORC data from Wittkowsky[2] and Rieder et al.[3])

Table S2—*[Section 1.3.1] Effect of CYP2C9 Genotype on Warfarin Dose Requirements*[4]

CYP2C9 Genotype	% Reduction in Warfarin Dose Requirement, Mean (95% CI)
*1/*1	Reference
*1/*2	19.6 (17.4-21.9)
*1/*3	33.7 (29.4-38.1)
*2/*2	36.0 (29.9-42.0)
*2/*3	56.7 (49.1-64.3)
*3/*3	78.1 (72.0-84.3)

Table S3—*[Section 2.2] Steady-State Maximum and Trough Plasma Concentration of Dabigatran After Daily or bid Doses of Dabigatran Etexilate Evaluated for the Prevention of VTE After Hip or Knee Replacement (150 mg and 220 mg od) and of Systemic Embolism in Nonvalvular Atrial Fibrillation (150 mg bid)*

Dose and Regimen	Cmax, ss, ng/mL	Ctrough, ss, ng/mL
150 mg od[a]	99	14
220 mg od[b]	183 (62-447)	37 (10-96)
150 mg bid[b]	184 (64-443)	90 (31-225)

Cmax = maximum plasma concentration; Ctrough = trough plasma concentration; od = once daily; ss = steady state. (Data from Eriksson et al.[5] and van Ryn et al.[6])

[a]Geometric mean. (From Eriksson et al.[5])

[b]Median with 5th and 95th percentiles. (From van Ryn et al.[6])

Table S4—*[Section 3.2] Maximum and Trough Plasma Concentration of Rivaroxaban in Patients Undergoing Total Hip Replacement After 5 d Rivaroxaban Dosing*

Dose and Regimen	Cmax, μg/L	Ctrough, μg/L
5 mg od	69.3	4.5
5 mg bid	39.8	8.4
10 mg od	124.6	9.1
10 mg bid	64.9	14.6
20 mg od	222.6	22.3
20 mg bid	141.9	35.1

See Table S3 legend for expansion of abbreviations. (Data from Mueck et al.[7])

 DOI: 10.1378/chest.11-2292

References

1. Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G; American College of Chest Physicians. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians evidence-based clinical practice guidelines (8th Edition). *Chest*. 2008;133(suppl 6): 160S-198S.
2. Wittkowsky AK. Pharmacology of warfarin and related anticoagulants. In: Ansell J, Oertel L, Wittkowsky A, eds. *Managing Oral Anticoagulation Therapy: Clinical and Operational Guidelines*. Vol. 1. St. Louis, MO: Facts and Comparisons; 2003: 29:1-29:15.
3. Rieder MJ, Reiner AP, Gage BF, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *N Engl J Med*. 2005;352(22):2285-2293.
4. Mentré F, Pousset F, Comets E, et al. Population pharmacokinetic-pharmacodynamic analysis of fluindione in patients. *Clin Pharmacol Ther*. 1998;63(1):64-78.
5. Eriksson BI, Dahl OE, Ahnfelt L, et al. Dose escalating safety study of a new oral direct thrombin inhibitor, dabigatran etexilate, in patients undergoing total hip replacement: BISTRO I. *J Thromb Haemost*. 2004;2(9):1573-1580.
6. van Ryn J, Stangier J, Haertter S, et al. Dabigatran etexilate—a novel, reversible, oral direct thrombin inhibitor: interpretation of coagulation assays and reversal of anticoagulant activity. *Thromb Haemost*. 2010;103(6):1116-1127.
7. Mueck W, Borris LC, Dahl OE, et al. Population pharmacokinetics and pharmacodynamics of once- and twice-daily rivaroxaban for the prevention of venous thromboembolism in patients undergoing total hip replacement. *Thromb Haemost*. 2008;100(3):453-461.

 DOI: 10.1378/chest.11-2292

THE AMERICAN JOURNAL ON ADDICTIONS

The American Journal on Addictions, 21: 257–262, 2012

ISSN: 1055-0496 print / 1521-0391 online
DOI: 10.1111/j.1521-0391.2012.00223.x

The Short Inventory of Problems—Modified for Drug Use (SIP-DU): Validity in a Primary Care Sample

Donald Allensworth-Davies, PhD, MSc,[1] Debbie M. Cheng, ScD,[2,3] Peter C. Smith, MD, MSc,[4] Jeffrey H. Samet, MD, MA, MPH,[3,5] Richard Saitz, MD, MPH[3,6]

[1]Data Coordinating Center, Boston University School of Public Health, Boston, Massachusetts
[2]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts
[3]Department of Medicine, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, Massachusetts
[4]Department of Medicine, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, Massachusetts
[5]Department of Community Health Sciences, Boston University School of Public Health, Boston, Massachusetts
[6]Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts

Primary care physicians can help drug-dependent patients mitigate adverse drug use consequences; instruments validated in primary care to measure these consequences would aid in this effort. This study evaluated the validity of the Short Inventory of Problems—Alcohol and Drugs modified for Drug Use (SIP-DU) among subjects recruited from a primary care clinic (n = 106). SIP-DU internal consistency was evaluated using Cronbach's alphas, convergent validity by correlating the total SIP-DU score with the DAST-10, and construct validity by analyzing the factor structure. The SIP-DU demonstrated high internal consistency (Cronbach's alpha for overall scale .95, subscales .72–.90) comparable with other SIP versions and correlated well with the DAST-10 (r = .70). Confirmatory factor analysis suggested an unacceptable fit of previously proposed factors; exploratory factor analyses suggested a single factor of drug use consequences. The SIP-DU offers primary care clinicians a valid and practical assessment tool for drug use consequences. (Am J Addict 2012;21:257–262)

INTRODUCTION

Primary care physicians are in a position to identify drug users years before they have medical complications or present for drug treatment. During the clinical encounter, primary care physicians can detect and then provide early intervention to help their patients become aware of and prevent adverse effects of drug use.[1] The individual consequences of drug use and dependence can be severe and may include loss of self-worth; loss of employment; loss of spouse, friends, and family; incarceration; as well as the development or worsening of chronic medical and/or psychiatric disorders. Understanding drug use consequences is central to addressing drug use and can provide motivation to the patient to address the problem. Understanding the severity of problems can also help a clinician determine what, if any, intervention is required (eg, a brief intervention in primary care for patients with milder symptoms or more intensive treatment for the dependent patient). Drug use instruments validated in a primary care setting are important; most such instruments were developed in specialty care settings and may not be appropriate for use in primary care.

One instrument that has been adapted to measure the consequences of alcohol and drug use combined is the Short Inventory of Problems (SIP). Originally developed and validated to measure alcohol consequences, the SIP has been modified to assess individual consequences related to substance (alcohol and other drug) use disorders (SIP-SUD), alcohol and other drug use (SIP-AD), drugs alone (SIP-D), and even bipolar disorder (SIP-BD).[2–4] None has been validated in primary care settings. The original SIP for alcohol is a 15-item instrument created by selecting representative items from the Drinker Inventory of Consequences (DrInC) to assess the self-attributed consequences of drinking in five domains (Physical, Intrapersonal, Social

Received March 11, 2011; revised March 31, 2011; accepted May 3, 2011.
Address correspondence to Dr. Allensworth-Davies, Data Coordinating Center, Boston University School of Public Health, 801 Massachusetts Avenue, Third Floor, Boston, MA 02118. E-mail: ddavies7@bu.edu.

Responsibility, Interpersonal, and Impulse Control) using two time frames (lifetime and recent consequences).[5] The Physical domain assesses acute and chronic physical states resulting from heavy drinking. The Intrapersonal domain asks about personal states of feeling associated with drinking such as guilt or shame. The Social Responsibility domain asks about consequences observable by others such as failing to do what was expected because of drinking. The Interpersonal domain assesses an individual's damage to or loss of personal relationships, including concern about drinking from family and friends. Finally, the Impulse Control domain asks about impulsive actions, risk-taking, increased use of other substances, and legal problems.[5,6]

The Inventory of Drug Use Consequences (InDUC) represented a next step in meeting the need for a standardized measure of both alcohol and other drug use consequences. The InDUC was created by taking the identical items used in the DrInC but modifying the wording from "drinking" to "drinking or using drugs." The InDUC is available in two versions: one that assesses lifetime consequences of substance use and another that assesses the frequency of recent problems (ie, past 3 months).[5] Like the DrInC, the InDUC was subsequently shortened to 15 items (of which 12 items are the same as those used in the SIP for alcohol) to enhance its clinical utility and became known as the SIP for alcohol or drugs (SIP-AD).[2]

It is this version of the SIP that was modified for this study to focus on other drug problems, and we hypothesized that it would have similar properties of internal consistency and validity as its recently developed cousin instrument, the SIP for alcohol modified for other drugs (SIP-D).[3] However, since the instrument used in this study is descended from a different branch of the SIP family tree than the SIP-D with three items that were not part of the SIP for alcohol, this instrument shall be referred to as the Short Inventory of Problems—Alcohol and Drugs modified for Drug Use (SIP-DU) to distinguish it from the SIP-D (Figure 1). Since no previous versions of the SIP have been validated exclusively in a primary care sample, the main objective of this study was to evaluate the validity of the SIP-DU in a sample of primary care patients.

MATERIALS AND METHODS

Participants

The study was conducted between October 2006 and June 2007 in the waiting room of a primary care clinic at an urban safety net hospital. To help minimize biased selection, subjects were systematically approached daily by one research associate according to a predetermined pattern based on waiting room seating, which was varied each day. Patients who were under the age of 18 were excluded, as were those who, in the judgment of the research associate, would be unable to complete the questionnaire because of limited English, cognitive impairment, or acute illness. People in the waiting room accompanying patients but who reported not themselves being patients of the clinic were also excluded. Eligible subjects were interviewed either before or after their primary care visit, or were scheduled to return at a later time for a research interview conducted in a private setting. All data were recorded anonymously without any unique identifiers.

Assessments

Subjects were first asked the single screening question validated in primary care to detect drug use, "How many times in the past year have you used an illegal drug or used a prescription medication for non-medical reasons?"[7] A cross-sectional assessment of alcohol and drug use was then conducted including the Drug Abuse Screening Test (DAST-10), a computerized version of the Composite International Diagnostic Interview (CIDI) Substance Abuse Module and the SIP-DU.[8,9] The questions included in the DAST-10 ask about possible involvement during the past 12 months with drugs by drug class, not including alcohol (eg, the use of prescribed or over-the-counter drugs in excess of the directions, any nonmedical use of drugs). As part of the CIDI, subjects were asked about past year use of illicit drugs (marijuana, cocaine, heroin, stimulants, or hallucinogens) and about past year non-medical use of prescription drugs. The SIP-DU asked questions about lifetime consequences of drug use (1 = Ever/0 = Never); subscale scores and the total score were then calculated by summing the item responses. All interviews were conducted by trained research staff (author PS trained the research associate by role play and by using the CIDI materials) in a private setting and data were recorded anonymously. All study procedures were reviewed and approved by the Institutional Review Board of Boston University Medical Center. Informed consent was given by all subjects before participating in the research.

Data Analysis

All analyses were conducted using SAS version 9.1 (SAS Inc., Cary, NC, USA). Scores were calculated for the total SIP-DU and each of the five subscales using the questions that asked about drug use consequences over the past 3 months. Cronbach's alpha (α) was used to assess internal consistency, or how closely related the set of questions are, and to evaluate how well the SIP-DU subscales measured distinct underlying constructs of drug use consequences. Internal consistency estimates were compared with published results from other SIP instruments.[4,5,10,11] The non-parametric Spearman's rho (ρ) was used to assess correlation of the SIP-DU with the DAST-10, a measure of drug involvement that is conceptually related (convergent validity). To evaluate whether the SIP-DU questions correspond to the previously proposed five concepts of Physical, Social, Interpersonal, Intrapersonal, and Impulse control consequences (construct validity) a confirmatory factor analysis was performed. The following criteria

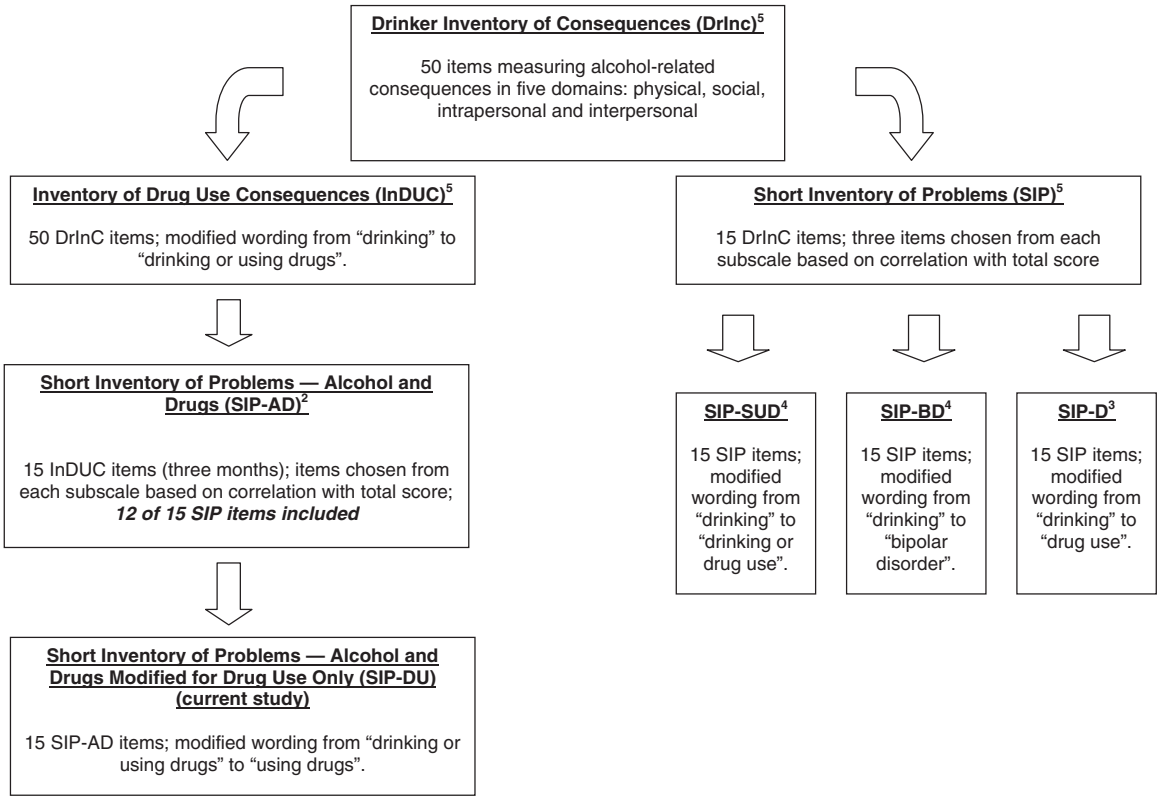


FIGURE 1. Derivation of the short inventories of problems. DrInC = Drinker Inventory of Consequences; InDUC = Inventory of Drug Use Consequences; SIP = Short Inventory of Problems; SIP-AD = Short Inventory of Problems—Alcohol and Drugs; SIP-SUD = Short Inventory of Problems—Substance Use Disorder; SIP-BD = Short Inventory of Problems—Bipolar Disorder; SIP-D = Short Inventory of Problems—Drugs; SIP-DU = Short Inventory of Problems—Drug Use.

were used to assess whether the model fit of the confirmatory factor analysis was acceptable: chi-square test (p-value .05); root mean square error of approximation (RMSEA) <.05; goodness of fit index (GFI) >.95; adjusted goodness of fit index (AGFI) >.90. If the analyses suggested an unacceptable model fit, an exploratory factor analysis was planned to further evaluate the possible underlying constructs of the SIP-DU. For the exploratory factor analysis, the number of factors to retain was determined based on the scree plot and the interpretability of the resulting factors. The exploratory factor analysis was performed using oblique rotation to allow for possible correlation among factors.

RESULTS

Sample Characteristics

Of 1,781 patients approached for screening 903 agreed to be screened and 394 were eligible for the study. Of the 394 eligible subjects, 303 arrived for the research interview and consented to participate in the study; 286 subjects completed the research interview. The sample for this analysis was limited ($n = 106$) to subjects who were identified as drug dependent based on their responses to the CIDI since they represent a relevant primary care population for assessment of drug use consequences (and the original SIP was derived from people with dependence). Study participants were predominantly male with a median age of 46 (range 22–74 years, Table 1). A majority was African American (58%), high school educated (73%), and reported English as their primary language (82%).

Scale Internal Consistency and Convergent Validity

Table 2 presents the scale internal consistency of the SIP-DU and, for comparison, those from the previously published DrInC and previous versions of the SIP, the SIP for alcohol use consequences, and the SIP for substance use disorders (SIP-SUD). These previous studies assessed versions of the SIP in several different populations including outpatients diagnosed with substance use and bipolar disorder, problem drinkers recruited from the community, and Emergency Department patients presenting with alcohol-related injuries.[4,5,10,11] Overall SIP-DU scale internal consistency was high (Cronbach's $\alpha = .95$) and the SIP-DU subscales had moderate-to-strong alphas (Physical = .71;

TABLE 1. Sample characteristics ($n = 106$)

Characteristic	
Female	37 (35)
Age	
Mean ± SD	46 ± 9
Median (range)	46 (22–74)
Education	
Some high school	35 (33)
High school graduate	41 (39)
Some college	25 (24)
College graduate or postgraduate education	5 (5)
Race	
American Indian/Alaskan Native	3 (3)
Black or African American	62 (59)
White	20 (19)
Unknown	21 (20)
Hispanic or Latino ethnicity	25 (24)
English is first language	87 (82)
Drug use in past year	
Marijuana or hashish	54 (51)
Cocaine	57 (54)
Prescription drugs without a prescription	32 (30)

Values are represented as *n* (%) except where otherwise noted; SD = standard deviation.

Social = .89; Interpersonal = .85; Intrapersonal = .86; Impulse = .82). Overall and across all five domains the SIP-DU internal consistency appeared as good as that previously reported for the related SIP measures by other studies. The total SIP-DU score had moderate-to-strong correlation with a conceptually related instrument, the DAST-10 (Spearman's $\rho = .71$), demonstrating convergent validity.

Construct Validity

Confirmatory factor analysis revealed that the five-factor structure did not provide an acceptable fit to the data based on each of the four fit indexes (chi-square $p < .0001$; RMSEA = .17; GFI = .74; AGFI = .58). An exploratory factor analysis was therefore conducted to further evaluate the possible underlying constructs of the SIP-DU. The scree plot suggested two underlying factors. In evaluating the two-factor model, each factor had multiple items with strong loadings and there was only one item that loaded on more than one factor. However, these factors were unable to be interpreted as distinct concepts of drug use consequences and did not correspond with the hypothesized subscales (Table 3). A three-factor model was then considered; however, this model demonstrated poor loadings on the third factor (data not shown). A single-factor model was then fit; all items demonstrated acceptable loadings (.598–.855) and the single-factor model explained 82% of the variance. Cronbach's alpha was .95 for the single-factor model suggesting a high degree of internal consistency. These results suggested a single-factor model provided the best fit to the data.

DISCUSSION

Recent initiatives to promote screening and brief intervention have resulted in improved screening tools for drug use and dependence in a variety of patient populations.[7,12,13] However, it is also important to assess the consequences of drug use once it is identified. Certainly reducing the negative physical, social, and personal consequences of drug use is an important quality of life issue for drug users,[14,15] and the involvement of primary care physicians in helping the patient find ways to reduce these consequences has the potential to improve both quality of care and outcomes. In fact, evidence suggests that higher quality primary care has the potential to reduce the odds of substance use.[16] The purpose of this study was to validate a version of the SIP that can be used in the primary care setting for drug use consequences. The SIP-DU and a single screening question about drug use provide time-efficient tools for primary care teams in identifying and caring for drug-using patients.[7]

Previous research has shown that the SIP and its variants (eg, the SIP-AD) are valuable measurement tools that can reliably assess negative consequences in patients with alcohol or drug use disorders.[2] Although one instrument that summarizes consequences of both alcohol and drug use can

TABLE 2. Internal consistency (Cronbach's alpha) of the SIP-DU compared to other SIP instruments

SIP scales	SIP-DU primary care ($n = 106$)	SIP-SUD ($n = 57$)	SIP community residents ($n = 153$)	SIP emergency department ($n = 404$)	DrInC ($n = 1,389$)
Total	.95	.93	.79	.95	.89
Physical	.72	.73	.64	.79	.67
Social	.90	.74	.62	.85	.76
Interpersonal	.85	.88	.61	.86	.76
Intrapersonal	.87	.76	.58	.89	.77
Impulse	.82	.62	.56	.73	.61

SIP = Short Inventory of Problems[10,11]; SIP-DU = Short Inventory of Problems—Drug Use; SIP-SUD = Short Inventory of Problems—Substance Use Disorder[4]; DrInC = Drinker Inventory of Consequences.[5]

TABLE 3. Exploratory factor analysis of the SIP-DU ($n = 106$)

Hypothesized subscale	SIP-DU questions	2-Factor solution		Single-factor solution
		Factor 1	Factor 2	Factor 1
Social	I have spent too much or lost a lot of money because of my drug use.	.939	−.045	.862
Social	I have failed to do what is expected of me because of my drug use.	.852	.036	.846
Physical	Because of my drug use, I have lost weight or not eaten properly.	.823	−.063	.736
Interpersonal	My family has been hurt by my drug use.	.848	.057	.861
Social	I have had money problems because of my drug use.	.754	.086	.795
Intrapersonal	I have been unhappy because of my drug use.	.731	.117	.799
Intrapersonal	I have lost interest in activities and hobbies because of my drug use.	.708	.122	.781
Intrapersonal	My drug use has gotten in the way of my growth as a person.	.602	.206	.748
Physical	My physical appearance has been harmed by my drug use.	.517	.311	.756
Impulse	When using drugs, I have done impulsive things that I regretted later.	−.129	.889	.620
Interpersonal	While using drugs I have said harsh or cruel things to someone.	−.013	.745	.612
Impulse	I have taken foolish risks when I have been using drugs.	.255	.639	.778
Interpersonal	A friendship or close relationship has been damaged by my drug use.	.305	.593	.789
Intrapersonal	When using drugs my personality has changed for the worse.	.347	.474	.729
Interpersonal	My drug use has damaged my social life, popularity or reputation.	.426	.464	.797
	Proportion of variance explained by factor	82%	9%	82%
	Eigenvalue	8.91	.95	8.91
	Standardized Cronbach's alpha	.91	.84	.95

SIP-DU = Short Inventory of Problems—Drug Use. Values in boxes indicate which items distinctly loaded onto either Factor 1 or Factor 2.

be useful as a global measure, tools that are more specific are needed for individual clinical care. Previous versions of the SIP have assessed consequences due to alcohol or drug use without distinguishing whether consequences were due to one or the other. This level of specificity is important when having a discussion with a patient about substance use consequences and when monitoring clinical progress of treatment outcomes in primary care. Having the ability to attribute consequences to alcohol and drugs separately may help clinicians and patients considering treatment options to prioritize and individualize interventions to mitigate the negative consequences experienced by the patient.

The results of this study suggest that the SIP-DU retains strong internal consistency when administered in primary care and that this is comparable to those reported in earlier studies.

The observed correlation of the SIP-DU with the DAST-10 is sufficiently high to suggest that the SIP-DU is measuring an independent construct of drug use consequences. This further highlights the potential usefulness of the SIP-DU in primary care settings with drug-using patients. This study also provides additional evidence that the original SIP can be successfully modified to separately assess consequences of drug use and alcohol use.

The unacceptable fit of the originally proposed five-factor structure for the SIP-DU is consistent with findings reported by Gillespie et al.[17] When they conducted confirmatory factor analysis with their data, they discovered that the loadings of the SIP-AD items were not consistent with a five-factor structure and concluded that a single-factor structure was most appropriate. The exploratory factor analysis findings of this study are consistent with those results.

This primary care sample was recruited from a single urban medical center, which may limit the generalizability of the findings. The patients at this medical center tend to be more socioeconomically disadvantaged, more racially and ethnically diverse, and may have more severe drug use or dependence than that seen at other hospitals that are not in urban underserved settings. Generalizability may also be limited to English-speaking patients willing to participate in a research study.

Since this was designed as a cross-sectional study, neither test-retest reliability nor the responsiveness of the SIP-DU to change over time could be evaluated. In addition, the sample size for this study was small and may not have been adequate to determine the true SIP-DU factor structure. Finally, unlike the full versions of the DrInC and InDUC, neither the original SIP for alcohol nor the SIP-DU contains a control subscale to detect careless responding.

As part of a growing interest in developing substance use measures relevant to primary care, the present study evaluated the validity of the SIP-DU in a primary care sample. As was recently found with the SIP-DU's cousin instrument, the SIP-D,[3] this analysis provides evidence of the SIP-DU's validity in measuring drug use consequences in a primary care setting. Consistent with previous research, a five-factor structure for the SIP-DU could not be confirmed by this study. Instead, the results suggested a possible single-factor structure for the SIP-DU. Future research might include additional psychometric analysis of the SIP-DU in other clinical settings and other geographical regions with larger sample sizes, and might also test use of the SIP-DU for identifying consequences of drug use among patients without dependence. Nevertheless, the SIP-DU offers clinicians a practical assessment tool for drug use consequences in a primary care setting.

Data collection and data analysis were supported by grant DA010019 from the National Institute on Drug Abuse, Bethesda, MD (Dr. Samet) and by grant AA010870 from the National Institute on Alcohol Abuse and Alcoholism, Rockville, MD (Dr. Saitz). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Drug Abuse, the National Institute on Alcohol Abuse and Alcoholism, or the National Institutes of Health.

We thank Dr. Sue Eisen for her review and commentary on the initial draft of this manuscript.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this paper.

REFERENCES

1. Haverkos HW, Stein MD. Identifying substance abuse in primary care. *Am Fam Physician.* 1995;52:2029–2035.
2. Blanchard KA, Morgenstern J, Morgan TJ, et al. Assessing consequences of substance use: Psychometric properties of the Inventory of Drug Use Consequences. *Psychol Addict Behav.* 2003;17:328–331.
3. Alterman AI, Cacciola JS, Ivey MA, et al. Reliability and validity of the alcohol short index of problems and a newly constructed drug short index of problems. *J Stud Alcohol Drugs.* 2009;70:304–307.
4. Bender RE, Griffin ML, Gallop RJ, et al. Assessing negative consequences in patients with substance use and bipolar disorders: Psychometric properties of the short inventory of problems (SIP). *Am J Addict.* 2007;16:503–509.
5. Miller WR, Tonigan JS, Longabaugh R. *The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse*. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 1995.
6. Forcehimes AA, Tonigan JS, Miller WR, et al. Psychometrics of the Drinker Inventory of Consequences (DrInC). *Addict Behav.* 2007;32:1699–1704.
7. Smith PC, Schmidt SM, Allensworth-Davies D, et al. A single-question screening test for drug use in primary care. *Arch Internal Med.* 2010;170:1155–1160.
8. Yudko E, Lozhkina O, Fouts A. A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. *J Subst Abuse Treat.* 2007;32:189–198.
9. Kessler RC, Abelson J, Demler O, et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). *Int J Methods Psychiatr Res.* 2004;13:122–139.
10. Feinn R, Tennen H, Kranzler HR. Psychometric properties of the short index of problems as a measure of recent alcohol-related problems. *Alcohol Clin Exp Res.* 2003;27:1436–1441.
11. Kenna GA, Longabaugh R, Gogineni A, et al. Can the short index of problems (SIP) be improved? Validity and reliability of the three-month SIP in an emergency department sample. *J Stud Alcohol.* 2005;66:433–437.
12. Amaro H, Reed E, Rowe E, et al. Brief screening and intervention for alcohol and drug use in a college student health clinic: Feasibility, implementation, and outcomes. *J Am Coll Health.* 2010;58:357–364.
13. D'Onofrio G, Degutis LC. Integrating project ASSERT: A screening, intervention, and referral to treatment program for unhealthy alcohol and drug use into an urban emergency department. *Acad Emerg Med.* 2010;17:903–911.
14. De Maeyer J, Vanderplasschen W, Broekaert E. Quality of life among opiate-dependent individuals: A review of the literature. *Int J Drug Policy.* 2010;21:364–380.
15. Stein MD, Mulvey KP, Plough A, et al. The functioning and well being of persons who seek treatment for drug and alcohol use. *J Subst Abuse.* 1998;10:75–84.
16. Kim TW, Samet JH, Cheng DM, et al. Primary care quality and addiction severity: A prospective cohort study. *Health Serv Res.* 2007;42:755–772.
17. Gillespie W, Holt JL, Blackwell RL. Measuring outcomes of alcohol, marijuana, and cocaine use among college students: A preliminary test of the Shortened Inventory of Problems—Alcohol and Drugs (SIP-AD). *J Drug Issues.* 2007;37:549–568.

The relationship between perceived discrimination and coronary artery obstruction

Brian J. Ayotte, PhD, [a] **Leslie R.M. Hausmann, PhD,** [b,c] **Jeff Whittle, MD, MPH,** [d,e] **and Nancy R. Kressin, PhD** [f,g]
North Dartmouth, and Boston, MA; Pittsburgh, PA; and Milwaukee, WI

Background Chronic stressors such as perceived discrimination might underlie race disparities in cardiovascular disease. This study focused on the relationship between perceived discrimination and risk of severe coronary obstruction while also accounting for multiple psychosocial variables and clinical factors.

Methods Data from 793 (629 white and 164 black) male veterans with positive nuclear imaging studies were analyzed. Participants were categorized as being at low/moderate or high risk for severe coronary obstruction based on results of their nuclear imaging studies. Hierarchical logistic regression models were tested separately for blacks and whites. The first step of the models included clinical factors. The second step included the psychosocial variables of optimism, religiosity, negative affect, and social support. The final step included perceived discrimination.

Results Perceived discrimination was positively related to risk of severe obstruction among blacks but not among whites after controlling for clinical and psychosocial variables. Similar results were found in patients who underwent coronary angiography (n = 311).

Conclusions Perceived discrimination was associated with risk of severe coronary obstruction among black male veterans and could be an important target for future interventions. (Am Heart J 2012;163:677-83.)

Cardiovascular disease affects disproportionately more blacks than whites in the United States.[1] However, the sources of this disparity are not clearly understood. One proposed mechanism underlying race differences in cardiovascular health is differential exposure to chronic stressors,[2] which includes racial discrimination.[3,4] As an unpredictable and uncontrollable stressor, perceived racial discrimination can elicit physiologic responses (eg, increased blood pressure) that can degrade health over time and result in cardiovascular disease over the long term.[5]

A growing body of literature suggests that perceived discrimination is related to diverse markers of cardiovascular disease. (Discrimination, while a real and objective dynamic, is most often measured by asking individuals their perceptions of whether/how discrimination occurred. We posit that such perceptions are the most important aspect to measure, because if discrimination occurs objectively but is not perceived by an individual, it may have less of an impact on health outcomes. For this reason, throughout the paper we will refer to 'perceived' discrimination, as that is the focus of this work.) Among blacks, but not whites, perceived racial discrimination was associated with higher levels of endothelin-1, which contributes to vascular remodeling, atherosclerotic plaques, renal dysfunction, and ultimately, hypertension and heart failure.[1] Chronic discrimination was also associated with coronary calcification, which is related to coronary obstruction, among a sample of black women.[6] In addition, perceived discrimination was positively associated with diastolic blood pressure reactivity and heart rate among black, but not white, women.[7] Although not all findings support the relationship between perceived discrimination and cardiovascular health,[8] there is significant evidence that perceived discrimination is related to cardiovascular health, particularly among black women.

A number of other psychosocial factors are associated with risk of cardiovascular disease. Negative affect, which includes a number of aversive mood states including disgust, nervousness, and anger,[9] is related to the development and expression of cardiovascular disease

From the [a]Department of Psychology, University of Massachusetts Dartmouth, North Dartmouth, MA, [b]Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System, Pittsburgh, PA, [c]Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, [d]Primary Care Division, Clement J. Zablocki VA Medical Center, Milwaukee, WI, [e]Department of Medicine, Medical College of Wisconsin, Milwaukee, WI, [f]VA Boston Healthcare System, Boston, MA, and [g]Section of General Internal Medicine, Boston University School of Medicine, Boston, MA.
Submitted October 7, 2010; accepted January 9, 2012.
Reprint requests: Brian J. Ayotte, PhD, Department of Psychology, University of Massachusetts Dartmouth, North Dartmouth, MA 02747.
E-mail: bayotte@umassd.edu
0002-8703/$ - see front matter

doi:10.1016/j.ahj.2012.01.006

and may be a mediator between perceived discrimination and cardiovascular disease.[10,11] Several protective psychological and social factors, including optimism, social support, and religiosity, have also been identified. *Optimism*, defined as the general expectation that good things will happen, is related to decreased mortality because of cardiovascular disease in black and white women.[12] Systematic reviews of the literature also suggest a consistent relationship between social isolation or low levels of social support and coronary artery disease.[13] Finally, religiosity is associated with positive health outcomes,[14] including the regression of coronary obstruction.[15] These factors have not been included in past studies examining the association between perceived discrimination and cardiovascular outcomes. Thus, it is unclear whether the influence of perceived discrimination on cardiovascular health persists after accounting for these other factors.

The objective of this study was to investigate the association of perceived discrimination with risk of severe coronary obstruction in a sample of white and black male veterans who had positive nuclear imaging studies. We hypothesized that increased levels of perceived discrimination would be related to greater risk of severe obstruction only among blacks even after controlling for clinical and behavioral risk factors (eg, smoking and high blood pressure) and psychosocial variables associated with cardiovascular health (ie, negative affect, optimism, social support, and religiosity). We also examined how the above variables were related to severity of coronary obstruction in the subset of veterans from our sample that underwent coronary angiography.

Methods

Study population

We used data from the Cardiac Decision Making Study, an observational cohort study of white and black veterans who had a cardiac nuclear imaging study performed between August 1999 and January 2001 at 1 of 5 Department of Veterans Affairs (VA) Medical Centers with on-site catheterization laboratories.[16] Overall, 2,335 of the 5,278 patients who were screened had a positive nuclear imaging study. Of these, 981 patients were excluded because they were unable to be contacted (n = 456), had a cardiac procedure in the preceding 6 months (n = 209), were not black or white (n = 102), had impaired mental status (n = 78), were in another study that would influence treatment of their coronary artery disease (n = 32), or other reasons (eg, impaired hearing, patient was a nonveteran, nuclear imaging study was performed for compensation and pension evaluation; n = 104). A total of 1,025 of the 1,354 eligible veterans agreed to participate, of which 793 (629 whites and 164 blacks) had complete data. We also collected coronary anatomy data in a subset of 311 patients (259 whites and 52 blacks) who had coronary angiography within 90 days of their imaging study.

Data

Sociodemographic information. Patients self-reported race, age (in years), and education (0, <12 years of education; 1, ≥12 years). Only patients who responded that they were white/Caucasian (coded as 0) or black/African American (coded as 1) were included in the study.

Clinical variables. Trained nurses abstracted clinical information, including diagnosis of hypertension and diabetes, current smoking status, prior revascularization, and prior myocardial infarction. These variables were coded as 1 if the conditions were present and 0 if they were not. We also collected data on a history of a diagnosis of hypercholesterolemia and if the participants were on antiplatelet or lipid-lowering therapy.

Perceived discrimination. Perceived discrimination was assessed with an adaptation of a previously published 7-item measure asking if the participant has ever been treated unfairly, been kept from doing something, or been made to feel inferior because of his or her race, ethnic group, or skin color in several domains, including (1) school, (2) getting a job, (3) at work, (4) getting medical care, (5) getting housing, (6) from the police or in the courts, or (7) on the street or in a public setting.[17] Because our patient population consisted of military veterans, we adapted the original measure by breaking the medical care question into "getting medical care at the VA" and "getting medical care outside the VA," and added "in the military" as an additional domain. The "yes" responses were counted to create a perceived discrimination score. Overall, the 9-item measure had good internal reliability among white (α = .75) and black (α = .88) participants. We also conducted the analyses using the original 7-item scale and found the same pattern of results, so only results using the 9-item scale are reported.

Psychosocial variables. *Negative affect* was assessed using a 9-item version[18] of the emotionality scale of the Eysenck Personality Inventory.[19] Participants responded "yes or "no" to items such as "Are you often uneasy, feeling that there is something you want without knowing it." The number of "yes" responses was counted to create a negative affect score. *Optimism* was assessed using a 10-item measure that asked participants to rate their agreement on items such as "In uncertain times, I usually expect the best" on a scale of 1 (strongly disagree) to 5 (strongly agree).[20] Responses were summed to create an optimism score. *Social support* was assessed using a validated 4-item scale,[21] which asked about satisfaction with family relationships, frequency of social contact with friends and relatives, and frequency of contact with someone the patient trusts and can confide in. Scores reflect the number of items to which the respondent answered "yes." *Religiosity* was assessed using the sum of 3 items asking about how often patients attended religious services; watched or listened to religious programs on TV or radio; or prayed, meditated, or studied the Bible or other religious text (0 = never/almost never, 7 = daily/more often).

Risk of severe coronary obstruction. We categorized risk of severe coronary obstruction using a modification of the methods of Bateman et al,[22] described in detail elsewhere.[23] Briefly, a board-certified general internist and a cardiology fellow classified the severity of each nuclear imaging study based on review of the official report, which did not include

Table I. Descriptive statistics

	Black (n = 164)			White (n = 629)			
Variable	**M**	**SD**	**Range**	**M**	**SD**	**Range**	***t***
Age (y)	60.60	10.81	37-85	63.24	9.46	31-84	3.08*
Perceived discrimination	3.27	2.90	0-9	0.41	1.08	0-8	–20.07*
Optimism	33.48	6.10	18-47	34.37	6.20	14-54	1.64
Negative affect	4.42	2.82	0-9	3.79	2.72	0-9	–2.66*
Religiosity	14.56	5.06	0-21	10.48	6.43	0-21	–7.54*
Social support	2.40	1.40	0-3	2.68	1.28	0-3	2.51*
	n	**%**		**n**	**%**		**χ^2**
Education							0.32
<12 y	49	29.9		174	27.7		
≥12 y	115	70.1		455	72.3		
Hypertension	138	84.1		486	77.3		3.67*
Diabetes	61	37.2		202	32.1		1.51
Prior revascularization	27	16.4		221	35.1		21.28*
Prior myocardial infarction	41	25.0		208	33.0		3.87*
Current smoker	54	32.9		188	29.9		5.30
Risk of coronary obstruction							0.03
Low/moderate	91	55.5		354	56.3		
Severe	73	44.5		275	43.7		

	Angiography results				
	Black (n = 52)		White (n = 259)		
	n	**%**	**n**	**%**	**χ^2**
Angiographic coronary obstruction					3.65
Mild/none	39	75.0	158	61.0	
Moderate/severe	13	25.0	101	39.0	

*$P < .05$

information on patient race. We considered patients with reversible lesions in the distribution of left anterior descending coronary artery or in both the right coronary artery and left circumflex artery to be at high risk, along with patients with increased lung uptake or transient ischemic dilation with exercise or pharmacologic stress. We considered patients with reversible lesions in just 1 of the right coronary artery or left circumflex artery to be at moderate risk and patients whose defects were very small or minimally reversible to be at low risk. For the current analysis, we collapsed patients into 2 categories; patients at moderate or low risk of severe coronary obstruction and those at high risk.

We also examined severity of coronary artery obstruction among participants who underwent coronary angiography. We classified coronary obstruction as severe if there was a stenosis of ≥70% in the left main coronary artery or all 3 major coronary systems. We classified nonsevere obstruction as moderate if there was at least 70% obstruction of the proximal left anterior descending artery. We defined mild obstruction as at least 1 coronary artery obstruction >70%, but not moderate or severe. We classified obstruction as none if there was no obstruction of >70%. For the current analysis, we collapsed obstruction into 2 categories: moderate/severe and mild/none.

Statistical analyses

We first examined race differences in all study variables using *t* tests or χ^2 tests. We then used hierarchical logistic regression models to examine the association between perceived discrimination and risk of severe coronary obstruction separately for blacks and whites. We chose to stratify the analyses by race because the experiences and consequences of racial discrimination are likely to be qualitatively different for blacks and whites, and the analytic strategies of including race as a covariate or attempting to adjust for variables confounded with race often can produce results that obscure racial differences.[24] Stratifying analyses by race is a recommended analytic technique that treats race as a "marker for differential experiences and exposures" (p. 303) and is less prone to mask important race differences.[25]

Each hierarchical model included 3 steps. Step 1 included sociodemographic and clinical characteristics that are known to differ by race or were associated with race in our sample (ie, age, education, hypertension, diabetes, prior revascularization, prior myocardial infarction, and current smoking status). We tested the impact of including additional clinical variables (receipt of antiplatelet or lipid lowering therapy, hypercholesterolemia), but they were not significant, so we omitted them from the final model. In the second step, we

Table II. Final logistic regression models predicting high (vs low/moderate) risk of severe coronary obstruction (nuclear imaging study results)

	Black (n = 164)			White (n = 629)		
		95%CI			95%CI	
Variable	**OR**	**Lower**	**Upper**	**OR**	**Lower**	**Upper**
Age	0.99	0.96	1.03	0.99	0.98	1.01
Education	0.92	0.78	1.09	0.99	0.93	1.05
Hypertension	1.49	0.55	4.03	1.27	0.85	1.90
Diabetes	1.19	0.56	2.56	0.79	0.55	1.12
Prior revascularization	2.02	0.77	5.32	1.07	0.73	1.54
Prior myocardial infarction	3.02*	1.24	7.33	1.65*	1.14	2.29
Current smoker	3.14*	1.29	7.90	1.46*	1.02	2.07
Social support	1.34	0.96	1.87	0.99	0.86	1.16
Optimism	0.93*	0.86	0.99	0.95*	0.95	0.98
Religiosity	0.96	0.90	1.04	0.97	0.95	1.01
Negative affect	1.04	0.88	1.23	0.91	0.84	1.99
Discrimination	1.27*	1.11	1.46	0.97	0.83	1.14

*$P < .05$

added several psychosocial variables that are related to cardiovascular health (ie, negative affect, optimism, social support, and religiosity). We added perceived discrimination in the final step. We repeated these steps for the analysis of coronary obstruction in the subsample of veterans who underwent coronary angiography.

The predictive power of the models was examined by use of the area under the receiver operating characteristic curve (ie, the *c* statistic). Values near 0.50 reflect a model with no apparent accuracy, values between 0.70 and 0.80 are considered good, values between 0.80 and 0.90 are considered excellent, and values above 0.90 are considered outstanding.[26]

This research was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (ECV 970-022 and RCS 02-066, NR Kressin, PI); the American Heart Association and the Pharmaceutical Roundtable (9970113N, NR Kressin, PI); Veterans Affairs Health Services Research and Development postdoctoral fellowship (BJ Ayotte); Veterans Affairs Health Services Research and Development Career Development Program (RCD 06-287, LRM Hausmann). The authors are solely responsible for the design and conduct of this study, all study analyses, and drafting and editing of the manuscript. The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Results

Descriptive statistics

Characteristics of the full sample are presented in Table I. Compared with blacks, whites were significantly older and less likely to be diagnosed with hypertension but were more likely to have had prior revascularization or prior myocardial infarction. Whites reported significantly less discrimination, negative affect, and religiosity, but significantly more social support than blacks. There were no race differences in education, optimism, and diabetes. Overall, 43.9% of participants were at high risk for severe coronary obstruction, based on their nuclear imaging studies; this percentage was similar in blacks and whites. Among the 311 patients who underwent coronary angiography, whites were significantly more likely than blacks to have moderate/severe coronary obstruction (39.0% vs 25.0%).

Logistic regressions

The final logistic regression models for the nuclear imaging study results are presented in Table II. Among blacks, step 1 (c_{model} = 0.66, χ^2_7 = 14.86, $P < .05$) indicated that prior myocardial infarction and smoking were positively associated with high (vs low/moderate) risk of severe coronary obstruction. In step 2 (c_{model} = 0.70, χ^2_4 = 10.36, $P < .05$), optimism was negatively related to risk of severe coronary obstruction. Finally, step 3 (c_{model} = 0.77, χ^2_1 = 13.45, $P < .05$) indicated that perceived discrimination was associated with an increase in risk of severe coronary obstruction after accounting for the variables in the first 2 steps of the model. Prior myocardial infarction, smoking, and optimism remained significant in the final model.

Among whites, smoking and myocardial infarction were related to high risk of severe coronary obstruction in step 1 ($c_{model} = 0.59$, $\chi^2_7 = 15.04$, $P < .05$). In step 2 ($c_{model} = 0.61$, $\chi^2_4 = 14.04$, $P < .05$), optimism was related to lower risk of coronary obstruction. Perceived discrimination was not related to risk of coronary obstruction among whites in step 3 ($c_{model} = 0.63$, $\chi^2_1 = 1.27$, $P > .05$).

The final logistic regression models for blacks and whites who underwent coronary angiography are presented in Table III. For blacks, step 1 ($c_{model} = 0.66$, $\chi^2_7 = 11.23$, $P < .05$) indicated that smoking was significantly related to the likelihood of moderate/severe coronary

Table III. Final logistic regression models predicting moderate/severe (vs mild/none) coronary obstruction (coronary angiogram)

	Black (n = 52)			White (n = 259)		
		95%CI			95%CI	
Variable	**OR**	**Lower**	**Upper**	**OR**	**Lower**	**Upper**
Age	0.99	0.88	1.12	1.00	0.98	1.03
Education	0.96	0.58	1.60	0.91	0.81	1.01
Hypertension	0.85	0.04	12.99	2.58*	1.15	4.75
Diabetes	1.07	0.27	5.62	1.02	0.55	1.78
Prior revascularization	1.45	0.79	3.72	0.38	0.46	1.58
Prior myocardial infarction	1.68	0.14	14.46	2.07*	1.12	3.83
Current smoker	1.49*	1.22	9.95	2.08*	1.17	3.72
Social support	1.20	0.35	4.08	1.11	0.80	1.53
Optimism	0.85*	0.73	0.99	0.97*	0.89	0.98
Religiosity	0.98	0.77	1.24	0.97	0.93	1.01
Negative affect	0.54	0.28	1.01	0.88	0.78	1.02
Discrimination	2.02	1.25	3.28	1.02	0.73	1.45

* $P < .05$

obstruction among blacks. In step 2 ($c_{model} = 0.79$, $\chi^2_4 = 13.46$, $P < .05$), optimism was related to a lower likelihood of moderate/severe obstruction. In step 3 ($c_{model} = 0.91$, $\chi^2_1 = 13.08$, $P < .05$), perceived discrimination was significantly related to increased likelihood of moderate/severe obstruction. Among whites, myocardial infarction and smoking were related to an increased likelihood of moderate/severe obstruction in step 1 ($c_{model} = 0.61$, $\chi^2_7 = 15.88$, $P < .05$). In step 2 ($c_{model} = 0.65$, $\chi^2_4 = 10.79$, $P < .05$), optimism was again related to a decreased likelihood of moderate/severe obstruction. Perceived discrimination was not associated with obstruction among whites in step 3 ($c_{model} = 0.68$, $\chi^2_1 = 1.01$, $P > .05$).

Discussion

In this study of black and white male veterans with abnormal nuclear imaging studies, we found that, among black men, greater perceptions of racial discrimination were related to increased risk for severe coronary obstruction and to angiographic coronary obstruction after controlling for clinical and psychosocial factors that are related to cardiovascular health. In addition, smoking, prior myocardial infarction, and optimism were related to nuclear imaging study results and coronary angiography results for blacks and whites.

These findings support and extend previous research linking perceived discrimination to indicators of cardiovascular disease[1,6,27,28] and are consistent with previous work that found these associations mostly among blacks.[1] Our results strengthen the evidence for a relationship between racial discrimination and cardiovascular disease in men, as previous work tended to focus on women. In addition, we controlled for psychosocial factors related to cardiovascular outcomes that have not been examined in previous research on the health correlates of perceived discrimination. Finally, by examining the association between perceived racial discrimination and coronary obstruction, we shed light on 1 possible causal pathway for the black-white disparity in cardiovascular deaths.

There are several possible explanations for the relationship we report. Chronic stress such as perceived discrimination is related to increased activation of the hypothalamic-pituitary-adrenocortical axis,[29,30] which results in increased production of cortisol. Cortisol, in turn, is independently related to risk of coronary artery disease and stenosis.[31,32] Research also suggests that perceived discrimination is related to hypertension[1] and increased cardiovascular reactivity,[7] which are both related to coronary obstruction.[33] In our results, perceived discrimination was associated with greater risk of severe coronary obstruction although we accounted for hypertension. Other possible contributors include increased endothelin-1 (a vasoconstrictor implicated in the development of vascular disorders[1,34]), e-selectin (an endothelial adhesion molecule associated with atherosclerosis[35]), and C-reactive protein (a marker of inflammation that is associated with cardiovascular disease[36]), all of which are related to perceived discrimination.

This study went beyond past studies of discrimination and cardiovascular disease indicators by including a number of psychosocial variables that are associated with cardiovascular health. Among blacks and whites, optimism was negatively related to the risk of severe coronary obstruction. Optimism might be related to cardiovascular health through factors such as hypertension,[37] as well as through other psychosocial factors (eg, increased social support, coping ability).[12]

This study found no significant associations of social support, negative affect, or religiosity with nuclear imaging study results or the results from coronary

angiography. One possible reason for these null findings is that our sample was restricted to persons with evidence for clinically significant coronary ischemia on nuclear imaging studies; these associations might be found in a sample that included persons without such evidence. It is also possible that these factors are not associated with coronary obstruction, per se, although they have been associated with other indicators of cardiovascular health. We chose to control for these variables in the current study, although it is possible that they might serve as moderators of the relationship between perceived discrimination and cardiovascular health because of their stress-buffering potential. We tested this possibility in supplemental analyses and found no evidence that optimism, negative affect, religiosity, or social support moderated the relationship between perceived racial discrimination and the nuclear imaging study results or the results from coronary angiography. With this being the first study to examine the relationships of these variables to nuclear imaging results, additional research is needed to determine whether the lack of significant relationships observed in the current study is seen in other populations.

These findings should be interpreted in the context of several limitations. First, our sample included only men with positive nuclear imaging study results—a population selected to be at high risk for coronary disease. This may explain why diabetes and hypertension, known risk factors for coronary artery disease, were not associated with the imaging results. Although physicians presumably included these clinical risk factors in their clinical decision making, it is unlikely that they incorporated perceptions of racial discrimination (or optimism). Thus, to the extent that perceived discrimination affects measurable risk factors for coronary artery disease, our estimates of the impact of perceived discrimination may be conservative. Future research should examine these relationships in people who are at low risk for coronary disease.

Second, our data were cross-sectional, thus precluding any causal interpretations. Third, we did not account for a number of clinical variables that are associated with risk of coronary obstruction (eg, various cholesterol subsets). Fourth, the sample size of black patients was relatively small, resulting in relatively large confidence intervals around our effect size estimates.

Finally, we assessed only racial discrimination, which was seldom a significant concern for white participants. This nonsignificant relationship between discrimination and risk of coronary obstruction should not suggest that discrimination is unimportant for whites. Previous studies indicate that nonracial discrimination is associated with cardiovascular outcomes, suggesting that it may be unfair treatment in general, rather than racial discrimination in particular, that leads to negative health effects.[38]

Despite these limitations, these results add to the growing literature on the relationship between perceived discrimination and cardiovascular health by investigating the relationship between perceived discrimination and risk of severe coronary artery obstruction. This is important because coronary artery disease is the most common form of heart disease and is the leading cause of death in the United States. Our findings underscore the possible negative health consequences of perceived discrimination among blacks. Future research should focus on identifying potential modifiable mediators of the relationship between perceived discrimination and cardiovascular disease so that the negative health consequences of perceived discrimination can be reduced.

References

1. Cooper R, Cutler J, Desvigne-Nickens P, et al. Trends and disparities in coronary heart disease, stroke and other cardiovascular diseases in the United States: findings of the National Conference on CVD Prevention. Circulation 2001;102:3137-47.
2. Baum A, Garofalo JP, Yali AM. Socioeconomic status and chronic stress: does stress account for SES effects on health? Ann N Y Acad Sci 1999;896:131-44.
3. Wyatt SB, Williams DR, Calvin R, et al. Racism and cardiovascular disease in African Americans. Am J Med Sci 2003;325: 315-31.
4. Smedley BD, Stith AY, Nelson AR, editors. Institute of Medicine, Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. Unequal treatment: confronting racial and ethnic disparities in health care. Washington: National Academies Press; 2002.
5. Clark R, Anderson NB, Clark VR, et al. Racism as a stressor for African Americans: a biopsychosocial model. Am Psychol 1999;54: 805-16.
6. Lewis TT, Everson-Rose SA, Powell LH, et al. Chronic exposure to everyday discrimination and coronary artery calcification in African-American women: the SWAN Heart Study. Psychosom Med 2006; 68:362-8.
7. Guyll M, Matthews KA, Bromberger JT. Discrimination and unfair treatment: Relationship to cardiovascular reactivity among African American and European American women. Health Psychol 2001;20: 315-25.
8. Albert MA, Ravenell J, Glynn RJ, et al. Cardiovascular risk indicators and perceived race/ethnic discrimination in the Dallas Heart Study. Am Heart J 2008;156:1103-9.
9. Watson D, Wiese D, Vaidya J, et al. The two general activation systems of affect: structural findings, evolutionary considerations, and psychobiological evidence. J Pers Soc Psychol 1999;76: 820-38.
10. Rozanski A, Blumenthal JA, Kaplan J. Impact of psychological factors on the pathogensis of cardiovascular disease and implications for therapy. Circulation 1999;99:2192-217.
11. Brondolo E, Brady N, Thompson S, et al. Perceived racism and negative affect: analyses of trait and state measures of affect in a community sample. J Soc Clin Psychol 2008;27:150-73.
12. Tindle HA, Chang YF, Kuller LH, et al. Optimism, cynical hostility, and incident coronary heart disease and mortality in the Women's Health Initiative. Circulation 2009;120:656-62.

13. Kuper H, Marmot M, Hemingway H. Systematic review of prospective cohort studies of psycosocial factors in the etiology and prognosis of coronary heart disease. Semin Vasc Med 2002;02:267-314.
14. Seeman TE, Dubin LF, Seeman M. Religiosity/spirituality and health: a critical review of the evidence for biological pathways. Am Psychol 2003;58:53-63.
15. Morris EL. The relationship of spirituality to coronary heart disease. Altern Ther Health Med 2001;7:96-8.
16. Kressin NR, Chang BH, Whittle J, et al. Racial differences in cardiac catheterization as a function of patients' beliefs. Am J Public Health 2004;94:2091-7.
17. Krieger N, Sidney S. Racial discrimination and blood pressure: the CARDIA Study of young black and white adults. Am J Public Health 1996;86:1370-8.
18. Floderus B. Psychosocial risk factors in relation to coronary heart disease and associated risk factors [monograph]. Norisk Hygienisk Tidskrift 1974(suppl 6):1.
19. Eysenck HJ, Eysenck SB. Manual of the Eysenck Personality Inventory. San Diego (Ca): Educational & Industrial Testing Service; 1968.
20. Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-master, and self-esteem): a reevaluation of the life orientation test. J Per Soc Psychol 1994;67: 1063-78.
21. Williams RB, Barefoot JC, Calif RM, et al. Prognostic importance of social and economic resources among medically treated patients with angiographically documented coronary artery disease. JAMA 1992; 267:520-4.
22. Bateman TM, O'Keefe Jr JH, Dong VM, et al. Coronary angiographic rates after stress single-photon emission computed tomographic scintigraphy. J Nucl Cardiol 1995;2:217-23.
23. Whittle J, Kressin NR, Peterson ED, et al. Racial differences in the prevalence of coronary obstructions among men with positive nuclear imaging studies. J Am Coll Cardiol 2006;47:2034-41.
24. Kaufman JS, Cooper RS. Commentary: consideration for use of racial/ethnic classification in etiological research. Am J Epidemiol 2001;154:291-8.
25. Jones CP. Invited commentary: "Race," racism, and the practice of epidemiology. Am J Epidemiol 2001;154:299-304.
26. Hosmer DW, Lemeshow S. Applied logistic regression. Hoboken (NJ): John Wiley & Sons, Inc; 2000.
27. Pascoe EA, Richman LS. Perceived discrimination and health: a meta-analytic review. Psychol Bull 2009;135:531-54.
28. Harris R, Tobias M, Jeffrys M, et al. Racism and health: the relationship between experience of racial discrimination and health in New Zealand. Soc Sci Med 2006;63:1428-41.
29. Rosmond R, Bjorntorp P. The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. J Intern Med 2000;247:188-97.
30. Fuller-Rowell TE, Doan SN, Eccles JS. Differential effects of perceived discrimination on the diurnal cortisol rhythm of African Americans and Whites. Psychoneuroendocrinology 2012;37: 107-18.
31. Smith GD, Ben-Shlomo Y, Beswick A, et al. Cortisol, testosterone, and coronary heart disease: Prospective evidence from the Caerphilly Study. Circulation 2005;112:332-40.
32. Koertge J, Al-Khalili F, Ahnve S, et al. Cortisol and vital exhaustion in relation to significant coronary artery stenosis in middle-aged women with acute coronary syndrome. Psychoneuroendocrinology 2002;27: 893-906.
33. Lovallo WR, Gerin W. Psycophysiological reactivity: mechanisms and pathways to cardiovascular disease. Psychosom Med 2003;65: 36-45.
34. Bohm F, Pernow J. The importance of endothelin-1 for vascular dysfunction in cardiovascular disease. Cardiovasc Res 2007;76: 8-18.
35. Friedman EM, Williams DR, Singer BH, et al. Chronic discrimination predicts higher circulating levels of E-selectin in a national sample: the MIDUS study. Brain Behav Immun 2009;23:684-92.
36. Lewis TT, Aiello AE, Leurgans S, et al. Self-reported experiences of everyday discrimination are associated with elevated C-reactive protein levels in older African-American adults. Brain Behav Immun 2010;24:438-43.
37. Raikkonen K, Matthews KA, Flory JD, et al. Effects of optimism, pessimism, and trait anxiety on ambulatory blood pressure and mood during everyday life. J Pers Soc Psychol 1999;76: 104-13.
38. Troxel WM, Matthews KA, Bromberger JT, et al. Chronic stress burden, discrimination, and subclinical carotid artery disease in African American and Caucasian women. Health Psychol 2003;22: 300-9.

Journal of Health Communication, 17:2–6, 2012
Copyright © Taylor & Francis Group, LLC
ISSN: 1081-0730 print/1087-0415 online
DOI: 10.1080/10810730.2012.717816

Introduction

Advancing Health Literacy Research

STACY COOPER BAILEY

Division of General Internal Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

LAUREN A. McCORMACK

Health Communications Program, RTI International, Research Triangle Park, North Carolina, USA

STEVEN R. RUSH

Health Literacy Innovations Program, United Health Group Minnetonka, Minnesota, USA

MICHAEL K. PAASCHE-ORLOW

Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

We are pleased to present this special issue on health literacy of the *Journal of Health Communication*. The journal continues to demonstrate a deep commitment to this important topic, offering a venue to exhibit leading health literacy research for the third year in a row. We hope that this special issue helps guide the overall direction of the field and serves as a springboard for lively discussion, innovative research, and practice initiatives.

As in the past, this special issue includes findings presented at the Health Literacy Annual Research Conference (HARC III), which was held on October 17–18, 2011 in Chicago, Illinois and was attended by more than 200 health literacy investigators. The 2011 conference was unique, as it was co-located with the International Conference for Communication in Healthcare (ICCH), an event dedicated to presenting cutting-edge research and educational strategies to improve how information and ideas are exchanged in health care settings—topics that are clearly related to the health literacy agenda. The coordination of these two conferences allowed researchers from countries around the world to share findings and build collaborations for the future. In doing so, it offered an exceptional opportunity for HARC to fulfill its primary objectives of (a) professional development, (b) advancing the science of health literacy

Address correspondence to Michael K. Paasche-Orlow, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA. E-mail: Michael.PaascheOrlow@bmc.org

research, and (c) promoting interdisciplinary research. We look forward to continuing our pursuit of these goals during HARC IV, which will be held on October 22–23, 2012 in Bethesda, Maryland.

The expansion of HARC, along with the growth in health literacy research, has led to the largest health literacy special issue yet. This issue includes over 20 articles and commentaries from across the United States and Europe, covering a number of key themes. Recognizing the growing diversity of the United States, multiple articles explore the links between health literacy, culture, and language proficiency. This includes a commentary from McKee and Paasche-Orlow calling for greater collaboration between researchers studying low health literacy and those focusing on limited English proficiency, as well as a conceptual framework by Lie, Carter-Pokras, Braun, and Coleman that integrates the concepts of health literacy and cultural competence for the purpose of improving education for health care professionals. An original research article from Shaw, Armin, Torres, Orzech, and Vivian explores the relationship between culture, health literacy, and chronic disease self-management among patients from four ethnic groups, while another from Sentell and Braun examines health status by health literacy and English proficiency among an ethnically and linguistically diverse study sample.

There has been increased recognition by researchers, clinicians, and government agencies of the importance of promoting patient-centered care and considering patient preferences in the design and implementation of interventions. Three articles in this issue respond to this movement by describing patients' preferences for receiving health information. Cawthon and colleagues discuss which components of their PILL-CVD intervention were considered beneficial according to patients, particularly those with low health literacy. Gaglio, Glasgow, and Bull report findings from an in-depth study on patients' preferred sources of health information. Finally, Guise, Koonce, Storrow, Kusnoor, and Ye reveal the importance of personalizing health education interventions by patients' preferred learning style and literacy skills.

As in the past, this special issue also includes articles focused on the conceptualization and measurement of health literacy. Squiers, Peinado, Berkman, Boudewyns, and McCormack introduce a conceptual framework for health literacy that builds on prior frameworks, but emphasizes a full pathway from the development of literacy skills to resulting health outcomes. Haun, Luther, Dodd, and Donaldson report on the conceptual and measurement variations between three commonly used measures of health literacy. Helitzer, Hollis, Sanders, and Roybal introduce the TALKDOC, an instrument designed to measure multiple health literacy competencies within the context of HPV and cervical cancer. Rubin presents an article on the construct of "listenability" (versus readability) as part of health literacy. Finally, Bann, McCormack, Berkman, and Squiers describe the validation of a shortened, 10-item version of their Health Literacy Skills Instrument. While development and validation activities for these tools should continue, next steps include incorporating these measures into future and ongoing research. This is essential to advance the field and ensure that validated, theory-based instruments are being used in practice.

This special issue expands upon the existing literature by including studies that further explore the association between health literacy and various health outcomes. Mitchell, Sadikova, Jack, and Paasche-Orlow assess the link between health literacy and re-hospitalization, while Mosher, Lund, Kripilani, and Kaboli examine the relationship between health literacy and medication-related outcomes, including

medication knowledge, adherence, and adverse drug events. Arnold and colleagues assess the role of health literacy in colorectal cancer screening knowledge, beliefs, and experiences. From an international perspective, Sahm, Wolf, Curtis, and McCarthy consider the prevalence and potential effects of limited health literacy in Ireland. Together, these articles serve to identify areas where additional research is needed; such findings have the potential to inform the design and implementation of future interventions.

Three articles describe efforts to improve the design, content, and efficacy of written materials from a health literacy perspective. Boxell and colleagues present information on the development and testing of a gynecological cancer information pamphlet in the United Kingdom, while King et al. evaluate the benefit of using symbols on prescription medication information leaflets. Liu and Rawl examine the role of text cohesion on the understandability of print materials, and Kaphingst et al. describe the development and testing of the Health Literacy INDEX, a tool that measures the health literacy demands of written health information materials. These articles highlight the ways in which written information can be enhanced, but also show the limitations of using print material alone to convey health information to patients.

Finally, a few articles in this special issue address topics that are of particular salience to health care organizations and providers. Weaver, Wray, Zellin, Gautam, and Jupka describe a process of assessing organizational health literacy in health centers serving vulnerable populations. Jager and Wynia examine how often physicians ask patients to "teach-back" information provided during clinical encounters; this article specifically examines the sociodemographic characteristics of patients who receive this health literacy via best practice versus those who do not. Finally, Bickmore and Paasche-Orlow discuss the potential impact of health information technology on health disparities and caution that greater inequalities will result without significant advances in health information technologies that are usable and accessible for people with low health literacy. The study by Gazmararian, Yang, Elon, Graham, and Parker explores this concern by examining the role of health literacy on women's ability to enroll electronically in the Text4Baby intervention.

Overall, this special issue covers a variety of topics, methods and perspectives. It takes a truly interdisciplinary approach, incorporating research from public health, medicine, nursing, anthropology, and communication science, among others. It includes perspectives of insurers, clinicians, patients, and researchers from the United States and Europe, using quantitative and qualitative methodologies to examine some of the most pressing questions in health literacy research. This issue acknowledges the limitations of a still relatively new field, exhibits a continued discussion about the conceptualization of health literacy and the use of imperfect measures—yet also shows great promise and progress. To move forward, we must continue to engage in this interdisciplinary dialogue, confronting and addressing our challenges, with the ultimate objective of advancing this important discourse.

References

Arnold, C. L., Rademaker, A., Bailey, S. C., Esparza, J. M., Reynolds, C., Liu, D., & Davis, T. D. (2012). Literacy barriers to colorectal cancer screening in community clinics. *Journal of Health Communication*, *17*(Suppl 3), 252–264.

Bann, C. M., McCormack, L. A., Berkman, N. D., & Squiers, L. B. (2012). The Health Literature Skills Instrument (HLSI): A 10-Item Short Form. *Journal of Health Communication*, *17*(Suppl 3), 191–202.

Bickmore, T. W., & Paasche-Orlow, M. K. (2012). The role of information technology in health literacy research. *Journal of Health Communication, 17*(Suppl 3), 23–29.

Boxell, E. M., Smith, S. G., Morris, M., Kummer, S., Rowlands, G., Waller, J., & Simon, A. (2012). Increasing awareness of gynecological cancer symptoms and reducing barriers to medical help-seeking: Does health literacy play a role? *Journal of Health Communication*, *17*(Suppl 3), 265–279.

Cawthon, C., Walia, S., Osborn, C. Y., Niesner, K. J., Schnipper, J. L., & Kripalani, S. (2012). Improving care transitions: The patient perspective. *Journal of Health Communication*, *17*(Suppl 3), 312–324.

Gaglio, B., Glasgow, R. E., & Bull, S. S. (2012). Do patient preferences for health information vary by health literacy or numeracy? A qualitative assessment. *Journal of Health Communication, 17*(Suppl 3), 109–121.

Gazmararian, J. A., Yang, B., Elon, L., Graham, M., & Parker, R. (2012). Successful enrollment in Text4Baby more likely with higher health literacy. *Journal of Health Communication*, *17*(Suppl 3), 303–311.

Giuse, N. B., Koonce, T. Y., Storrow, A. B., Kusnoor, S. V., & Ye, F. (2012). Using health literacy and learning style preferences to optimize the delivery of health information. *Journal of Health Communication*, *17*(Suppl 3), 122–140.

Haun, J., Luther, S., Dodd, V., & Donaldson, P. (2012). Measurement variations among brief health literacy instruments: Implications for research and practice. *Journal of Health Communication*, *17*(Suppl 3), 141–159.

Helitzer, D., Hollis, C., Sanders, M., & Roybal, S. (2012). Assessing the "other" health literacy competencies—Knowledge, dispositions, and oral/aural communication: Development of TALKDOC, an intervention assessment tool. *Journal of Health Communication*, *17*(Suppl 3), 160–175.

Jager, A. J., & Wynia, M. K. (2012). Who gets a teach-back? Patient-reported incidence of experiencing a teach-back. *Journal of Health Communication*, *17*(Suppl 3), 294–302.

Kaphingst, K. A., Kreuter, M. W., Casey, C., Leme, L., Thompson, T., Cheng, M.-R., … Lapka, C. (2012). Health literacy INDEX: Development, reliability, and validity of a new tool for evaluating the health literacy demands of health information materials. *Journal of Health Communication*, *17*(Suppl 3), 203–221.

King, S. R., McCaffrey, D. J. III, Bentley, J. P., Bouldin, A., Hallam, J., & Wilkin, N. E. (2012). The influence of symbols on the short-term recall of pharmacy-generated prescription medication information in a low health literate sample. *Journal of Health Communication*, *17*(Suppl 3), 280–293.

Lie, D., Carter-Pokras, O., Braun, B., & Coleman, C. (2012). What do health literacy and cultural competence have in common? Calling for a collaborative health professional pedagogy. *Journal of Health Communication*, *17*(Suppl 3), 13–22.

Liu, C.-J., & Rawl, S. M. (2012). Effects of text cohesion on comprehension and retention of colorectal cancer screening information: A preliminary study. *Journal of Health Communication*, *17*(Suppl 3), 222–240.

McKee, M. M., & Paasche-Orlow, M. K. (2012). Health literacy and the disenfranchised: The importance of collaboration between limited English proficiency and health literacy researchers. *Journal of Health Communication*, *17*(Suppl 3), 7–12.

Mitchell, S. E., Sadikova, E., Jack, B. W., & Paasche-Orlow, M. K. (2012). Health literacy and 30-day post-discharge hospital utilization. *Journal of Health Communication*, *17*(Suppl 3), 325–338.

Mosher, H. J., Lund, B. C., Kripalani, S., & Kaboli, P. J. (2012). Association of health literacy with medication knowledge, adherence, and adverse drug events among elderly veterans. *Journal of Health Communication*, *17*(Suppl 3), 241–251.

Rubin, D. L. (2012). Listenability as a tool for advancing health literacy. *Journal of Health Communication*, *17*(Suppl 3), 176–190.

Sahm, L. J., Wolf, M. S., Curtis, L. M., & McCarthy, S. (2012). Prevalence of limited health literacy among Irish adults. *Journal of Health Communication*, *17*(Suppl 3), 100–108.

Sentell, T., & Braun, K. (2012). Low health literacy, limited English proficiency, and health status in Asians, Latinos, and other racial/ethnic groups in California. *Journal of Health Communication*, *17*(Suppl 3), 82–99.

Shaw, S. J., Armin, J., Torres, C. H., Orzech, K. M., & Vivian, J. (2012). Chronic disease self-management and health literacy in four ethnic groups. *Journal of Health Communication*, *17*(Suppl 3), 67–81.

Squiers, L., Peinado, S., Berkman, N., Boudewyns, V., & McCormack, L. (2012). The health literacy skills conceptual framework. *Journal of Health Communication*, *17*(Suppl 3), 30–54.

Weaver, N. L., Wray, R. J., Zellin, S., & Gautam, K., & Jupka, K. (2012). Advancing organizational health literacy in health care organizations serving high needs populations: A case study. Journal of Health Communication, *17*(Suppl 3), 55–66.

The Breast Journal

ORIGINAL ARTICLE

The Addition of Internists to a Breast Health Program

Tracy A. Battaglia, MD, MPH,*,†,‡ Mary Beth Howard, BS,* Maureen Kavanah, MD,† Marianne N. Prout, MD, MPH,‡ Chava Chapman, MBBch, MPH,*,† Michele M. David, MD, MBA, MPH,*,† Renee McKinney, MD,*,† Andrea Kronman, MD, MSc,*,† Tara Dumont, MD,*,† and Karen M. Freund, MD, MPH*,†,‡

*Department of Medicine, Section of General Internal Medicine, Women's Health Unit, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine; †Belkin Breast Center, Boston Medical Center; and ‡Department of Epidemiology, Boston University School of Public Health, Boston, MA

■ **Abstract:** With the increases in complexity of care for breast health concerns, there is a growing need for efficient and effective clinical evaluation, especially for vulnerable populations at risk for poor outcomes. The Breast Health Center at Boston Medical Center is a multidisciplinary program, with internists providing care alongside breast surgeons, radiologists, and patient navigators. Using a triage system previously shown to have high provider and patient satisfaction, and the ability to provide timely care, patients are assigned to either a breast surgeon or internist. From 2007 to 2009, internists cared for 2,408 women, representing half of all referrals. Women served were diverse in terms of race (33% black, 30% Hispanic, 5% Asian), language (34% require language interpreter), and insurance status (51% had no insurance or public insurance). Most presented with an abnormal screen (breast examination 54% or imaging 4%) while the remainder were seen for symptoms such as pain (26%), non-bloody nipple discharge (4%), or risk assessment (7%). A majority of final diagnoses were made through clinical evaluation alone (n = 1,760, 73%), without the need for additional diagnostic imaging or tissue sampling; 9% (n = 214) received a benign diagnosis with the aid of breast imaging; 19% (n = 463) required tissue sampling. Only 4% went on to see a breast surgeon. Internists diagnosed 15 incident cancers with a median time to diagnosis of 19 days. Patient and provider satisfaction was high. These data suggest that a group of appropriately trained internists can provide quality breast care to a vulnerable population in a multidisciplinary setting. Replication of this model requires the availability of more clinical training programs for non-surgical providers. ■

Key Words: breast, diagnostic, health services, prevention, screening

Increased attention to the early detection of breast cancer has resulted in more women participating in breast cancer screening (1). Concurrently, new and changing screening guidelines, imaging options, risk assessment methodologies, preventive strategies, and diagnostic algorithms and techniques have made the early detection of breast cancer increasingly complex. As available screening methods are not highly specific, abnormal findings are common and often require clinical expertise, such that the demand for specialists in breast health has increased (2,3). Yet, the number of breast cancers diagnosed as a result is quite low; the average positive predictive value (proportion of women with abnormal screening exams who have breast cancer) is approximately 2.0% for women below age 50, and 4.7% for women over age 50.

The complexity of breast care can lead to failures in the process of care, which lead to lapses or discontinuation of breast health care that have the potential to impact survival for those found to have a malignancy (4,5). Lack of access to timely appointments with breast care subspecialists has been shown to contribute to such delays (6). The delivery of these services to underserved populations introduces additional complexity as a result of the additional burden of well-documented socio-cultural barriers to accessing care (7–9). Indeed, delays in diagnosis after an abnormal screening test for these underserved populations,

Address correspondence and reprint requests to: Tracy A. Battaglia, MD, MPH, Associate Professor of Medicine Women's Health Unit, Section of General Internal Medicine, Evans Department of Medicine, Boston Medical Center and Women's Health Interdisciplinary Research Center, Boston University School of Medicine. 801 Massachusetts Ave, Suite 470, Boston, MA 02118, or e-mail: Tracy.Battaglia@bmc.org.

DOI: 10.1111/j.1524-4741.2011.01179.x

including those from racial/ethnic minority and low socioeconomic groups, have been found to contribute to worse breast cancer outcomes (10–13).

A number of societies have promoted guidelines to address the management of women with breast problems (14–17). While many of these guidelines suggest that diagnostic evaluation requires expertise of surgical oncologists or breast surgeons, there is a lack of evidence that this model is required for quality care (18). As many women in need of breast health consultation do not have surgical needs, alternative care models are needed to improve access for those who do. Therefore, we describe here a new care model for breast health and breast concerns that incorporates general internists with an expertise in women's health. This clinical practice, previously shown to improve adherence to timely follow-up of screening abnormalities (19), provides services to a disproportionate share of vulnerable and traditionally underserved women. We provide a description of the program structure and function along with evaluation metrics as evidence of its impact on quality care.

Origin

The Breast Health Program at Boston Medical Center (the teaching hospital of Boston University School of Medicine) was created in 1989 through collaboration between the Women's Health Unit (within the Department of Internal Medicine), Surgical and Medical Oncology, and Breast Imaging. It was developed out of the recognition that diagnostic evaluation for breast disease requires knowledge and skills beyond the level of traditionally trained internists, yet surgical providers were inundated with non-surgical referrals. As a result, breast surgeons had fewer appointments available for those truly in need of their service.

Philosophy

The Breast Health Program at Boston Medical Center is a multidisciplinary and interdisciplinary diagnostic evaluation program. The goal of the program is to provide efficient and comprehensive diagnostic services and prevent delay in cancer diagnosis and care. Our multidisciplinary team includes seven internists, three breast surgeons, 2–4 radiologists specializing in breast imaging, and three pathologists with expertise in cytopathology and/or surgical pathology. Interdisciplinary support is available with nurses, genetic counselors, social workers, patient navigators, medical assistants, and breast imaging technicians. Providers aim to deliver comprehensive services at each visit to minimize potential gaps in care, while addressing a wide range of breast conditions and barriers to care as they present in the clinical setting.

Features

Communication Communication between providers is facilitated by several features (1). Space: When the practice began, there was no allocated space for all providers, so patient evaluation and care were accomplished through a "virtual center." With growth and documented success of the program (19,20), space for all clinical consultation and breast imaging was created in a dedicated facility to allow for real time communication and improve opportunities for formal and informal consultation on cases (2). Schedules: Schedules are coordinated to ensure that both an internist and breast surgeon staff the center during each available clinical session, at which time a full spectrum of screening and diagnostic breast imaging services are also available (3). Regular conferences: A multidisciplinary breast conference, which includes representation from internal medicine as well as all oncology specialists (surgical, medical, and radiation oncology, pathology, and clinical trials), meets weekly to review all cases with ambiguous or inconclusive cytopathology as well as all newly diagnosed cancer cases to create comprehensive diagnostic and treatment plans. Special seminars are also organized quarterly to discuss challenges in management of benign breast disease, such as protocols for use of screening MRI among high-risk populations (4). Meetings: Leadership from internal medicine, surgery, and radiology meet monthly with nursing and administrative staff, to review organizational and management issues along with quality metrics. Additional interdisciplinary monthly meetings include IT staff, internal medicine, program coordinator, and patient navigators to review quality indicators and refine reporting systems and clinical protocols.

Triage System The Breast Health Program developed a central intake system to triage referrals to either a surgeon or internist based on patient risk of cancer. The aim was to ensure that women most likely to need surgical intervention are initially seen by a breast surgeon (Fig. 1). Triage decisions are made by an experienced practice triage nurse, and

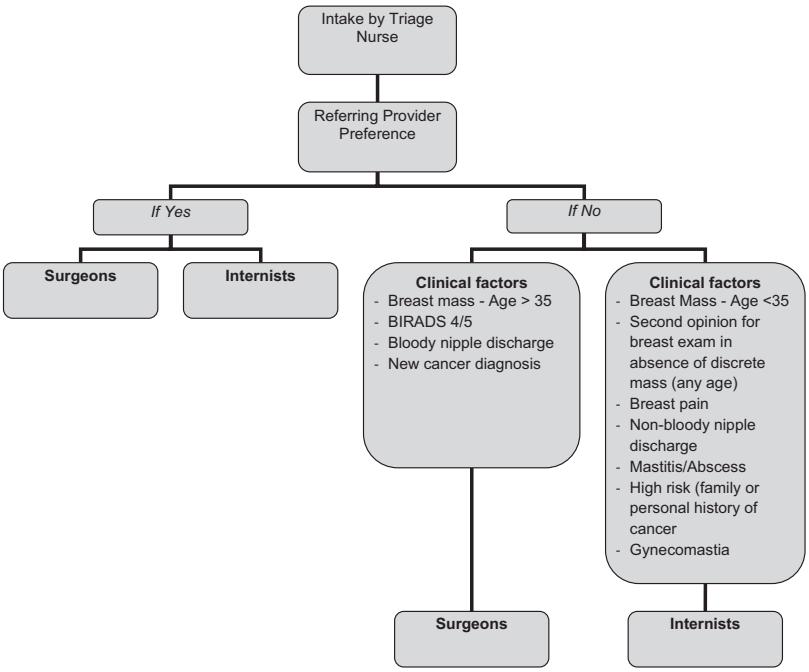


Figure 1. Breast health program triage system.

result in an equal distribution of patient referrals being scheduled with internists and surgeons. We have previously shown that this triage system minimizes referral error, decreases time to surgical appointment for those with a malignancy, and satisfies referring providers (20). A preclinical review includes assessment of potential need for diagnostic imaging or other services at the initial or follow-up visit. Interdisciplinary staff will consult with providers to schedule this additional diagnostic imaging to coincide with the patient visit.

Patient Navigation Breast Health Patient Navigators have provided care management services to help our vulnerable patients access needed services with funding from the Avon Foundation since 2001 (21). Two full-time patient navigators identify women referred to the practice, conduct outreach to identify barriers to care, work with providers to develop an individual action plan to overcome those barriers, and utilize the electronic medical record to track all women until diagnostic resolution. Their daily activities are documented in a log within patient charts in the Electronic Medical Record (EMR). Patient navigators are nonclinical personnel usually with bachelor or master's level training, who have experience caring for inner-city, low income patients and/or are from one of the communities served by the medical center.

Tracking Prospective tracking by navigators ensures that patients are followed through completion of all diagnostic workup and treatment as recommended by their provider. Once patients have been referred into our practice, we utilize the hospital EMR to retrieve detailed patient information, including demographics, appointment scheduling, navigator outreach, visits with providers and diagnostic procedures. This information is consolidated into one automated data tracking system, which provides patient navigators with a list of active patients in need of specific services, such as additional imaging (mammograms, MRIs or ultrasound), pending cytology or pathology. Navigators work directly with the internists to ensure that timely resolution is achieved. Navigator supervisors review tracking logs weekly to ensure resolution of outstanding cases. This database also provides quality improvement evaluation data to guide our program management (19,22).

Services Provided by Internists

Internists provide a wide range of services, including (a) counseling on evidence for additional surveillance, breast cancer risk assessment, and discussion of genetic testing, fertility planning, and preventive strategies, including prophylactic surgery, and chemoprevention; (b) clinical evaluation of abnormal breast findings such as nipple discharge or breast masses; and (c) management of symptomatic benign disease,

including breast pain, cysts, acute and recurrent mastitis/abscess. Internists include Fine Needle Aspiration biopsies (FNAB) and Core biopsies into their clinical competencies for two reasons: (a) to facilitate rapid diagnosis should a cancer be present; and (b) to complete diagnostic evaluation on patients not requiring surgical intervention.

Training of Internal Medicine Physicians Most internists, including those with a focus on women's health, have insufficient training and skills to conduct a comprehensive breast health assessment (23). We developed a training system, which pairs trainees with an experienced internal medicine breast provider in the clinical setting for 4–8 hours weekly over a 3–4 month period to provide consulting in a supervised setting. These trainees, all of whom have previously completed internal medicine residency training, also shadow surgical providers. Trainees review all breast imaging with radiologists to develop skills in correlating clinical findings with diagnostic imaging. Trainees practice FNAB and core needle biopsies (with the support of pathologists and cytopathologists) and must demonstrate proficiency in these procedures before they perform them independently. For the first year of practice, trainees are paired with each of the practicing internists on a rotating basis in an effort to increase exposure to multiple practice styles.

EVALUATION METHODS

We report on data collected from the EMR for all patients seeking care from January 1, 2007 through December 31, 2009. Patient demographics are retrieved from the electronic scheduling system (SDK©; Esker Inc., Madison, WI), while all clinical information is retrieved from the medical charts in EMR (General Electric (GE); Centricity^{EMRTM}, https://www2.gehealthcare.com/portal/site/usen/menuitem.e8b305b80b84c1b4d6354a1074c84130/?vgnextoid=5bb454fbded30210VgnVCM10000024dd1403RCRD&vgnextfmt=default&productid=4bb454fbded30210VgnVCM10000024dd1403). Reason for referral to the breast practice is documented in the EMR by the Triage Nurse at the time of intake, along with the referral source. Final diagnosis was determined from pathology and cytology reports, from diagnostic imaging if no pathology or cytology was performed, and from ICD-9 billing diagnoses when no additional imaging or pathology was performed. To ensure that no cancer diagnoses were missed, all patients seen were cross referenced with the Boston Medical Center Cancer registry. Patient satisfaction was continually measured using the institution-wide Press Ganey survey (Press Ganey Associates, South Bend, Indiana), a commercial survey sent to a random subset of all patients with a response rate that represents 20% of the population served. IRB approval was obtained for all data collected.

RESULTS

Patient Population

Patient volume for the Breast Health Program has increased from 651 unique patients in 2001 to 901 unique patients in 2009 as a result of the expansion of internists into the practice. From January 1, 2007 through December 31, 2009, 3,110 patient visits

Table 1. Patient Population *n* = 2,408 seen by Internists, Boston Medical Center Breast Health Program January 2007–December 2009

	n (%)
Age	
<18	56 (2)
19–39	1,520 (63)
40–64	723 (30)
>65	109 (5)
Race/ethnicity	
Black	810 (33)
Hispanic	715 (30)
White	555 (23)
Asian	111 (5)
Other	217 (9)
Language	
English	1,635 (68)
Spanish	446 (19)
Haitian Creole	149 (6)
Other*	215 (9)
Insurance	
Private	1,169 (49)
Public	824 (34)
None†	415 (17)
Referring location	
Neighborhood health centers	1,424 (59)
Medical center-based referral	721 (30)
Outside hospital/private practice	263 (11)
Referral reason	
Abnormal exam	1,302 (54)
Abnormal imaging	92 (4)
Nipple discharge	89 (4)
Pain	620 (26)
Family history of breast cancer	151 (6)
Other	154 (6)

*Other category includes: Vietnamese (2%), Portuguese (2%), Portuguese Creole (2%), all other language categories <2%.
†Includes free care, no insurances, and insurance provider unknown.

occurred for a total of 2,408 unique patients seen by medical internists. As described in Table 1, the populations served are diverse and reflective of the urban population seeking care at an inner-city safety-net institution. As would be expected from the triage protocol (see Fig. 1), the majority of women seen by internists were under age 65 (95%), and over half between 19 and 39 years of age (63%). A majority of women were of minority race or ethnicity (33% Black, 30% Hispanic, 5% Asian); while 23% were non-Hispanic White. Thirty-four percent required a language interpreter during their visit, and over half (51%) had publicly funded or no health insurance at the time of their referral.

Table 1 also shows the referral network for those patients seen by internists. About 60% of referrals come from a network of over 15 affiliated community health centers, and about 30% from practices within the Boston Medical Center, including the emergency department, obstetrics and gynecology, family medicine, primary care internal medicine, and geriatrics. The remaining are referred from outside providers, with few self-referrals. The two most frequent reasons for referral triaged to internists are abnormal breast exam (54%) and breast pain (26%), followed by family history of breast cancer, abnormal imaging, and nipple discharge (Table 1).

Services Rendered and Final Diagnoses

Table 2 shows that a majority of final diagnoses were made through clinical evaluation alone (n = 1,760, 73%), without the need for additional diagnostic imaging or tissue sampling. Nine percent (n = 214) received a benign diagnosis with the aid of breast imaging, most commonly from ultrasonography, followed by diagnostic mammography and rarely Magnetic Resonance Imaging. The remaining 19% of final diagnoses (n = 463) were made as result of tissue sampling, obtained mainly via percutaneous tissue sampling, while only 4% went on to see a breast surgeon for further evaluation. Of those referred to surgeons, the median age was 33 years (range 15–74 years) and the most common reason for referral was an abnormal clinical breast exam (55%), followed by abnormal imaging and patient preference for surgical excision of a benign process. A total of 15 women never returned to complete recommended care after an initial evaluation. For many, there was evidence that they transferred care to other institutions.

From 2007 to 2009, a total of 15 breast cancers (Table 3) were diagnosed among all the patients initially referred to internists, versus 502 cancers diagnosed by an initial breast surgeon evaluation during the same time period. Median age of these 15 women was 55 years (range 32–76) and most were Non-White (n = 10). The most common presenting symptom was an abnormal examination (n = 6), which ranged from asymmetry to a discrete mass. The majority of women (n = 9) were being followed by internists for long-term breast health care in the setting of either elevated cancer risk, pain, or in the context of an established primary care relationship with the internist, which explains why some were not ''triaged'' directly to a surgeon according the existing triage protocol. As a result, most women (n = 9) obtained a percutaneous tissue diagnosis prior to

Table 2. Final Diagnoses for Patients Referred to Internists, Boston Medical Center Breast Health Program January 2007–December 2009[*]

Final diagnosis	Clinical diagnoses (ICD-9 codes) *n* (%)	Radiologic diagnoses (imaging reports) *n* (%)	Tissue diagnoses (cytology or pathology reports) *n* (%)	Total frequency *n* (%)
Cancer	0	0	15 (3)	15 (0.6)
Atypical ductal hyperplasia	0	0	3 (1)	3 (0.1)
Benign breast diagnoses[†]	200 (12)	0	336 (73)	536 (22)
Cyst/fibrocystic changes	675 (39)	0	109 (23)	784 (33)
Breast Pain	568 (30)	0	0	568 (23)
Benign, decided after abnormal initial imaging	100 (6)	214 (100)	0	314 (13)
Family history/high risk	63 (4)	0	0	63 (3)
Infection/inflammation of the breast	23 (1)	0	0	23 (1)
Benign nipple discharge	32 (2)	0	0	32 (1)
Musculoskeletal issue of the chest wall	13 (0.8)	0	0	13 (0.5)
Dermatologic issue of the breast	42 (2)	0	0	42 (2)
Lost to follow-up	0	0	0	15 (0.6)
Total	1,716	214	463	2,408

[*]Final diagnoses reported for patients seen by Internists only.
[†]Includes fibroadenoma, papilloma, lipoma, physiologic nodularity.

Table 3. Description of Incident Cancer Cases at Boston Medical Center Breast Health Program January 2007–December 2009

Patient characteristic	*n* = 15 (%)
Median age at diagnosis (range)	55 (32–76)
Race/ethnicity	
Black	7 (46)
White	5 (33)
Asian	1 (7)
Hispanic	1 (7)
Other	1 (7)
Presenting symptoms	
Abnormal clinical breast exam	6 (40)
Breast pain	2 (13)
Abnormal imaging	3 (20)
Risk assessment	3 (20)
Breast abscess	1 (7)
Median time to diagnosis (range)*	19 days (3, 69)
Stage at diagnosis	
0	4 (27)
1	4 (27)
2	2 (13)
3	3 (20)
Missing†	2 (13)

*Measured as time from date of referral to program to date of pathologic diagnosis.
†Patients went to outside institutions for cancer treatment.

being referred to a surgeon. The median time from initial referral to cancer diagnosis was 19 days (range 3–69 days). For one outlier, 69 days elapsed from referral to cancer diagnosis; during that time period, there are 15 documented attempts by the patient navigators to schedule for follow-up care, along with evidence of other active comorbid medical conditions. Most women were Stage 0 or 1 at the time of diagnosis (n = 8), while 5 did present with Stage 2 or 3 disease. Cross referencing with the hospital cancer registry revealed no additional cancer diagnoses on patients seen by internists during this time period.

Satisfaction with Care

Press Ganey data collected by Boston Medical Center indicates that over 90% of patients report being satisfied or very satisfied with access to care and confidence in provider. This figure remained constant over the 3 years. Referring provider satisfaction with the program has previously been shown to be high (20). Our surgeons have found this to be a very useful model in that it allows them to take care of patients more likely to be surgical and avoid backing up their practices with non-surgical cases.

DISCUSSION: IMPACT OF OUR MODEL

We present comprehensive process and outcomes data for a model of care that delivers diagnostic services by including a group of dedicated internists focused on breast care working collaboratively alongside breast surgeons. Our data demonstrate that appropriately trained internists can provide timely, quality evaluation using a team approach with surgical specialists and an evidence-based triage protocol. Therefore, we believe that this model of care for breast health is worthy of adoption at other sites.

While many models of care for management of breast masses state the need for a surgical oncology evaluation, no data support this need (24). To the contrary, several of our findings support that the use of non-surgical providers may be both reasonable and appropriate. First, we have shown here and in our prior work (20), using an evidence-based triage protocol results in successful risk stratification such that internists are evaluating a population with an inherently low probability of malignancy. Furthermore, even among the 3–4% with a cancer diagnosis, there were no delays in diagnosis that may be attributed to the system and none of the nearly 3,000 women served went on to a breast cancer diagnosis over the ensuing few years to suggest any missed diagnoses. It is also worth noting that the majority of patients were given a final diagnosis with only a clinical diagnosis, while only 19% required cytopathologic or tissue sampling, most of which was performed prior to a surgical consultation. With the support of pathologists, Internists are trained to perform these non-invasive procedures. Finally, this model has potential to improve access to breast surgeons by high-risk patients, thus preventing delays in surgery. This potential impact is supported by the average time to surgical evaluation at our institution of 2 days. The high rates of Stage 2 and 3 cancers among those diagnosed by internists are reflective of the overall rates of advanced cancer seen among the population of mostly minority, low-income patients served at our institution, and not linked to any delays in access to diagnostic services.

The American Society of Breast Disease recently developed a framework of quality indicators related to multidisciplinary and interdisciplinary care (25). These quality indicators include the same outcome measures our program monitors, including early diagnosis, decreased time to visit for most at risk, patient satisfaction, and provider satisfaction. The ability of our program to deliver quality indicators recommended by leading authorities is further evidence that it is worthy of replication in other settings.

While internists are not traditionally trained or prepared to conduct diagnostic breast health evaluation,

our findings support their ability to provide quality care. Our patient population is reflective of the vulnerable patients served at other safety net intuitions who are at high risk for presenting with late-stage disease. Our diverse population presented with unusual manifestations of common and rare diseases, including diabetic mastopathy, mammary tuberculosis, breast sarcoidosis, cat scratch disease, abscesses, and recurrent nonlactating mastitis that often require complex medical management. Having a broad background in pathophysiology and disease management is critical for these rare and/or comorbid conditions. Lastly, this group can provide surveillance to high-risk women, for prevention and screening options—areas in which internists are well trained.

There are limitations to our evaluation metrics. First, we do not have data prior to implementation of this model to allow for an historical comparison. Also, we do not have cost or cost effectiveness data. Future research should address the financial implications of similar models, as it is likely that improving timely access for all patients, especially those with cancer, along with the provision of comprehensive services in one visit with an internist, is likely to have financial benefits to the health care system.

In summary, a model of specifically training internists to provide comprehensive breast health services within a multidisciplinary breast cancer center is a novel and effective model that results in high-quality care that is worthy of adoption when serving similar vulnerable populations. More clinical breast health training programs are needed for non-surgical providers to replicate this type of multidisciplinary program.

Acknowledgments

We are grateful for the contributions and ongoing support of Dr Michael Stone, Chief, Surgical Oncology, at Boston Medical Center.

REFERENCES

1. Kilmer G, Roberts H, Hughes E, *et al.* Surveillance of certain health behaviors and conditions among states and selected local areas—Behavioral Risk Factor Surveillance System (BRFSS), United States, 2006. *MMWR Surveill Summ* 2008;57:1–188.

2. Anderson LM, May DS. Has the use of cervical, breast, and colorectal cancer screening increased in the United States? *Am J Public Health* 1995;85:840–2.

3. Zapka JG, Stoddard AM, Costanza ME, Greene HL. Breast cancer screening by mammography: utilization and associated factors. *Am J Public Health* 1989;79:1499–502.

4. Weingart S, Saadeh M, Simchowitz B, *et al.* Process of care failures in breast cancer diagnosis. *J Gen Intern Med* 2009;24:702–9.

5. Richards MA, Westcombe AM, Love SB, Littlejohns P, Ramirez AJ. Influence of delay on survival in patients with breast cancer: a systematic review. *Lancet* 1999;353:1119–26.

6. Hung WK, Chan SW, Suen DT, Chan MC, Lau Y, Yip AW. Triaging referral to a specialist breast clinic. *ANZ J Surg* 2006;76:310–2.

7. Harper AP. Mammography utilization in the poor and medically underserved. *Cancer* 1993;72(Suppl 4):1478–82.

8. Panel PsC. *Report of the Chairman 2000–2001. Voices of a Broken System: Real People, Real Problems*. Bethesda: National Cancer Institute: National Institutes of Health, 2003.

9. Wolff M, Bates T, Beck B, Young S, Ahmed S, Maurana C. Cancer prevention in underserved African American communities: barriers and effective strategies—a review of the literature. *WMJ* 2003;102:36–40.

10. Ries L, Eisner M, Kosary C. *SEER Cancer Statistics Review, 1975–2000*. Bethesda, MD: National Cancer Institute, 2003.

11. Jones BA, Dailey A, Calvocoressi L, *et al.* Inadequate follow-up of abnormal screening mammograms: findings from the race differences in screening mammography process study (United States). *Cancer Causes Control* 2005;16:809–21.

12. Chang SW, Kerlikowske K, Nápoles-Springer A, Posner SF, Sickles EA, Pérez-Stable EJ. Racial differences in timeliness of follow-up after abnormal screening mammography. *Cancer* 1996;78:1395–402.

13. Caplan LS, May DS, Richardson LC. Time to diagnosis and treatment of breast cancer: results from the National Breast and Cervical Cancer Early Detection Program, 1991–1995. *Am J Public Health*. 2000;90:130–4.

14. American College of Surgeons. National Accreditation Program for Breast Centers. Available at: <http://accreditedbreastcenters.org/> (accessed November 30, 2009) Chicago.

15. Blamey RW, Cataliotti L. EUSOMA accreditation of breast units. *Eur J Cancer* 2006;42:1331–7.

16. Anderson BO, Carlson RW, Kaufman CS, Kiel KD. Ensuring optimal interdisciplinary breast care in the United States. *Breast J* 2009;15:569–70.

17. National Consortium of Breast Centers, Inc. National Quality Measures for Breast Centers. Available at: <http://www.nqmbc.org/default.htm> (accessed December 1, 2009) Warsaw, Indiana, 2004.

18. Link J. History and overview of comprehensive interdisciplinary breast centers. *Surg Oncol Clin N Am* 2000;9:147–57.

19. Battaglia TA, Roloff K, Posner MA, Freund KM. Improving follow-up to abnormal breast cancer screening in an urban population. *Cancer* 2007;109(Suppl 2):359–67.

20. Stone MD, Norton S, Mendez JE, Hirsch E. Positive impact of a breast-health triaging system on breast-care access and physician satisfaction. *Am J Surg* 2007;194:482–7.

21. Longest BB, Young G. Coordination and communication. In: Shortell SK, K A, eds. *Health Care Management: Organization Design and Behavior*, 4th edn. Albany, NY: Delmar Publishers, 2000.

22. Wells KJ, Battaglia TA, Dudley DJ, *et al.* Patient navigation: state of the art or is it science? *Cancer* 2008;113:1999–2010.

23. Fletcher S. Following up abnormal breast cancer screening results: lessons for primary care clinicians. *J Am Board Fam Pract* 2000;13:152–4.

24. Lopez A, Barker G, Bhattacharryya A, *et al.* Expedited breast care: a new model in breast health. *Digital Mammography* 2008;5116:515–21.

25. Ensuring Optimal Interdisciplinary Breast Care in the United States. *ABSD Colloquium*. Dallas, TX: American Society of Breast Disease, 2006.

J Ambulatory Care Manage
Vol. 35, No. 1, pp. 38–49

Feasibility of Chronic Disease Patient Navigation in an Urban Primary Care Practice

Tracy A. Battaglia, MD, MPH; Lois McCloskey, MPH, DrMPH; Sarah E. Caron, MPH; Samantha S. Murrell, MPH; Edward Bernstein, MD; Ariel Childs, MPH; Hannah Jong, MPH; Kelly Walker, MPH; Judith Bernstein, PhD, RNC

Abstract: The purpose of this study was to evaluate the feasibility of incorporating chronic disease navigation using lay health care workers trained in motivational interviewing (MI) into an existing mammography navigation program. Primary-care patient navigators implemented MI-based telephone conversations around mammography, smoking, depression, and obesity. We conducted a small-scale demonstration, using mixed methods to assess patient outcomes and provider satisfaction. One hundred nine patients participated. Ninety-four percent scheduled and 73% completed a mammography appointment. Seventy-one percent agreed to schedule a primary care appointment and 54% completed that appointment. Patients and providers responded positively. Incorporating telephone-based chronic disease navigation supported by MI into existing disease-specific navigation is efficacious and acceptable to those enrolled. **Key words:** *chronic disease prevention, mammography, motivational interviewing, patient navigation, primary care, screening, vulnerable populations*

Author Affiliations: *Women's Health Unit, Section of General Internal Medicine, Evans Department of Medicine and Women's Health Interdisciplinary Research Center (Dr Battaglia and Mss Caron and Murrell), Emergency Medicine (Dr E. Bernstein), Boston University School of Medicine, and Community Health Sciences Department, Boston University School of Public Health (Drs McCloskey and J. Bernstein, Mss Childs, Jong, and Walker), Boston, Massachusetts.*

This work was partially supported by the Evans Center for Interdisciplinary Biomedical Research Junior Faculty Merit Award 2010-2011 at Boston University (http://www.bumc.bu.edu/evanscenteribr/) a pilot grant from the Department of Medicine and the BU Bridge (Boston University Clinical and Translational Science Institute UL1RR025771) and the Avon Foundation Safety Net Grant. The authors thank the patient navigators who worked so hard to make this research possible, as well as Caitlyn Barthelmes for her assistance in the conduct of the study intervention.

The authors have disclosed that they have no significant relationships with, or financial interest in, any commercial companies pertaining to this article.

Correspondence*: Tracy A. Battaglia, MD, MPH, Women's Health Unit, Section of General Internal Medicine, Evans Department of Medicine and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, MA 02118 (Tracy.Battaglia@bmc.org).*

DOI: 10.1097/JAC.0b013e31822cbd7c

IN RESPONSE TO the fragmented and complex nature of cancer care, patient navigation was designed to help vulnerable patients overcome barriers to accessing cancer care services through the support of a lay health care worker connected to the clinical team (Dohan & Schrag, 2005; Freeman et al., 1995). In 2007, a study conducted by the Health Resources and Services Administration identified "navigator" as one of the 5 prevailing models for community health work and 18% of employers reported patient navigation as a specific strategy they use to conduct community health work (Health Resources and Services Administration, 2007). Navigators, now a standard of care in some practice settings (American College of Surgeons, 2009), assist patients with scheduling, transportation,

insurance, childcare, translation, and any other issues that interfere with timely access to health care. Successful navigation programs traditionally target a specific disease such as breast cancer (Battaglia et al., 2007; Freeman et al., 1995) without taking into account competing medical comorbidities, patient health priorities, or patient preferences. The majority of programs target single-site cancers (Bastani et al., 2010; Battaglia et al., 2007; Burhansstipanov et al., 2010; Chen et al., 2008; Ell et al., 2007; Percac-Lima et al., 2009) in specialty practice settings, while few target multiple cancer sites in a primary care setting (Dietrich et al., 2006; Ford et al., 2006; Guadagnolo et al., 2010).

More recently, the patient-centered medical home (PCMH or "medical home") is emerging as a promising health care delivery model that will help transform the US health care system to improve quality of care delivered and reduce costs (Stange et al., 2010). The principles of the medical home (American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, American Osteopathic Association, 2007) build on the primary care model by incorporating the evidence-based chronic care model along with the use of office systems and health information technology (Scholle et al., 2010). This model of care is also patient-centered with an emphasis on ensuring that patients get the care they need where and when they want it. Patient navigation, as a care coordination process, fits well within many of the domains of the medical home, including care coordination, access to care, communication, and tracking and registry. Researchers are beginning to assess the feasibility and acceptability of chronic disease and primary care navigation (Ferrante et al., 2010; Lubetkin et al., 2010). However, little is known about the feasibility of expanding the role of disease-specific navigation to target chronic disease navigation.

Motivational interviewing (MI) is a form of patient-centered counseling directly tailored to enhance readiness to change behavior. It was originally developed for use by psychologists in the context of substance abuse treatment (Bernstein et al., 2005, 2007, 2009a, 2009b; Hettema et al., 2005; Kaner et al., 2007; Stein et al., 2009; Whitlock et al., 2004) but has been widely adapted to address screening and intervention (both primary and secondary preventions) for other behaviors that affect health status (Brug et al., 2007; Kressin et al., 2009; Mason, 2009; Ockene et al., 1999). Motivational interviewing is based on the principle that exploring and resolving ambivalence about taking action is the central motivational task that prepares people for change.

The objective of this study was to evaluate the feasibility of expanding a primary care breast navigation program to address additional medical comorbidities and allow patients to set their own priorities for action through the use of brief motivational interview skills (B-MI). We hypothesize that incorporating multiple chronic disease screenings into navigation outreach is feasible and patients will respond positively. We believe that incorporating MI, with its emphasis on patient choice, into the current navigator model will facilitate agenda-setting and behavior change.

METHODS

This feasibility study (Bowen et al., 2009) was designed as a clinical quality improvement initiative and thus did not have a continuous control group. We conducted a small-scale demonstration study, using mixed methods to determine the (1) practicality (ability to carry out the intervention), (2) limited efficacy (intended effects on clinical care), and (3) acceptability (satisfaction, intent to use) of telephone-based chronic disease navigation in an urban safety net primary care practice. All participants received the B-MI intervention March-July 2010, and we used electronic medical record (EMR) chart abstraction, navigator logs, and qualitative interviews, respectively, to measure outcomes. The Boston Medical Center institutional review board approved this study.

Theoretical framework

Brief motivational interview is an adaptation of standard motivational techniques to the time pressures of the clinical visit

and delivery by both clinicians and outreach workers. Motivational interviewing teaches health care providers to follow a specific algorithm for a conversation about decreasing risk factors for disease, enhancing patients' and providers' self-efficacy for behavioral change (Ockene et al., 1996), and facilitating referral for mental health and other treatment if indicated (Bernstein et al., 2005). The literature (Apodaca & Longabaugh, 2009) suggests that effect on outcomes is directly related to the amount of change talk produced during such conversations. When individuals believe that they will be successful (effective) at making desired changes in their behavior, it is more likely that they will actually be able to do so (Bandura, 1977). The role of the patient navigator in this study of chronic disease navigation is to use B-MI skills via telephone contact to foster self-efficacy, thus increasing the probability of keeping a mammogram appointment and enhance the likelihood of making effective contact with a primary care physician to discuss other prevention issues such as weight management, depression risk, and smoking.

The importance of patients' participation in their own medical care and decision making is well documented, particularly for those with chronic illness (Greenfield et al., 1985, 1988; Kaplan et al., 1989). A key element in a participatory style of patient care is soliciting the patient's priorities and preferences during the medical visit (Kaplan et al., 1995; Marvel et al., 1999). Thus, our MI intervention also included prompts for patients to identify their priorities and set the agenda for their future primary care visit.

Study population

This study was conducted in Boston at an urban safety net hospital serving low-income, racial/ethnic minority residents. Female patients of the internal medicine practice aged 51-70 years with a documented primary care visit in the past 24 months were eligible if they had at least 1 prior mammogram and more than 18 months had elapsed since their most recent documented mammogram. Because of limited resources, non-English speakers were excluded. In addition, patients were excluded if their primary care provider deemed them ineligible on the basis of the presence of comorbidities that impaired their ability to participate, including a current diagnosis of breast cancer or travel out of the country during the study period.

Navigator training

Five experienced primary care–based breast health patient navigators participated in this study. In addition to clinical navigation experience, each completed certification for the core competencies of a community health care worker offered by a regional education center. Each completed an intensive 2-week training workshop led by the Brief Negotiated Interview-Active Referral to Treatment Institute (BNI-ART) (http://www.ed.bmc.org/sbirt/). The training covered basic clinical information about obesity/weight management, smoking cessation, and depression; how to screen patients for weight, smoking, and depression; and how to apply motivational intervention techniques to screenings. Navigators engaged in role-playing and other experiential learning activities. A booster training was held 1 month after study enrollment began.

Intervention protocol

Navigators contacted eligible patients by phone. Multiple attempts to contact patients were made at varying times and days of the week. If contact was not made after 6 attempts, the patient was deemed "unreachable." If contact was made, verbal consent to participate was obtained. The navigator then conducted the B-MI intervention. The conversation included discussions on smoking status, depression, obesity, and participant health priorities, as well as mammography screening. Participants were asked whether weight was a concern for them and whether they were currently a smoker and were screened for depression using the Patient Health Questionnaire-2 (Kroenke et al., 2003; Lowe et al., 2005). Navigators also asked participants to list their top health priorities. Navigators worked with each participant to schedule mammography and primary care appointments and recorded the reason if participants were unwilling to schedule

appointments. Navigators offered to accompany participants to appointments. Reminder calls were made by navigators prior to scheduled appointments to discuss any barriers to attendance and to work with participants to overcome those barriers.

Data sources and data collection

Practicality data elements were measured using navigator logs. Navigators documented their activities, including telephone attempts and participant responses on encounter log forms.

Efficacy data and participant characteristics were collected from manual chart abstraction using the EMRs. Research assistants followed a standard operating procedure to obtain participant sociodemographic information, body mass index, current smoking status, and mammogram and primary care appointment information. Clinical data were abstracted up to 30 days after scheduled appointment dates for each participant.

Acceptability data were collected using semistructured qualitative one-on-one interviews with key informants: a subset of participants (n = 11), all navigators (n = 5), and a subset of primary care providers (n = 5). Participants with a scheduled primary care appointment were eligible for the qualitative portion of this study, were consented, and then contacted by a research assistant by telephone to conduct the interviews. These interviews included a 2-part phone call—the first after the initial call with the navigator and the second after the scheduled primary care appointment. Participants were asked about their experiences and the nature of the conversation with the navigator. Likert scales (range 1–5, 5 = most helpful) were incorporated to assess the participant's satisfaction with components of navigation.

The navigator satisfaction interviews were conducted in person and assessed their satisfaction as a motivational interviewer and with providing navigation for other health issues in addition to the usual breast cancer screening. Provider interviews were administered via telephone and assessed the physicians' satisfaction with the role that the navigator had in preparing their patients for the primary care visit.

All interviews were digitally recorded and transcribed. The research team read the transcripts multiple times and, using an iterative process, developed a coding scheme to reflect the key concepts raised in the narratives. The team coded each transcript independently, using HyperRESEARCH software (ResearchWare, Inc, Randolph, Massachusetts). The team then met to discuss the coded narratives and the major themes and subthemes that emerged.

Measures

The following chronic disease screenings were defined as positive:

1. Overweight/obese = a body mass index $\geq$ 25 recorded in the EMR.
2. Current smoker = a positive response to the navigator asking, "Do you smoke?"
3. Depression = score of 3 or more on the navigator administered Patient Health Questionnaire–2 (range: 0-6).

Participant-identified health priorities were elicited by the navigator.

Practicality outcome measures include (1) eligibility rate of female primary care patients, (2) rate of eligible patients successfully reached by telephone outreach, and (3) rate of consent among those reached. Primary *efficacy* outcome measures include (1) completion of a mammography appointment and (2) completion of a primary care appointment. Completion was defined as attending an appointment within 30 days of the original appointment scheduled by the navigator. *Acceptability* outcome measures include (1) participant and provider satisfaction, (2) participant willingness to have navigators accompany them to appointments, and (3) participant willingness to speak with primary care provider about their health priorities.

Statistical analysis

Descriptive statistics were performed using SAS version 9.1. Basic frequencies provided practicality and efficacy outcomes. Chi-square and *t* tests were used to examine associations between enrollment status and sociodemographic characteristics.

RESULTS

Figure 1 displays *practicality* outcomes. An electronic query of the EMR generated a list of 671 active female general internal medicine patients aged 51 to 70 years whose last documented mammogram was 18 or more months prior. These patient charts were reviewed by research staff for enrollment eligibility. Of 671 patients, only 270 (40%) were considered eligible. The most common reason for ineligibility was non-English language (135/401, 39%), while the remainder did not meet the prespecified eligibility criteria, such as age or time since last mammogram. Among those eligible for enrollment, almost a third (87/270, 32%) could not be reached via telephone outreach. When eligible patients were reached, the majority (109/270, 68%) agreed to participate.

The Table shows the sociodemographic characteristics of the eligible cohort reached via telephone. More than half of all participants were blacks, on public health insurance, and most had a documented visit with their primary care doctor within the previous 12 months. Rates of overweight/obese were substantial. Compared with those who declined to participate, those enrolled into the study were younger, more likely to be on public insurance, more likely to be current smokers, and less likely to have had a primary care visit in the past year.

The Table also shows prevalence of chronic disease for the enrolled group. The most prevalent condition was overweight/obesity (25% and 55%, respectively), and about one-third of participants screened positive for both smoking and/or depression. Overall, 93% of participants screened positive for one of the three chronic disease states; 39% had more than 1 condition (33% had 2 conditions, and 6% had all 3 conditions).

Overall, 67% of participants reported personal health priorities when asked. The majority of these participants (85%) reported health priorities that were congruent with the screening program priorities (36% weight/obesity, 10% depression, and 9% smoking). Weight was the most common patient-reported health priority. Diabetes (4%) was the most common patient-reported disease-specific health priority that was not included as a program priority. Symptoms such as pain or discomfort accounted for 13% of reported health priorities. Most (83%) stated that they would be willing to talk to their primary care providers about their health priorities.

Figure 2 shows the clinical efficacy of the pilot intervention. Overall, 93% of participants scheduled a mammogram appointment as a result of a navigator encounter, and after 30 days, 73% completed their mammogram test. Meanwhile, 71% of participants

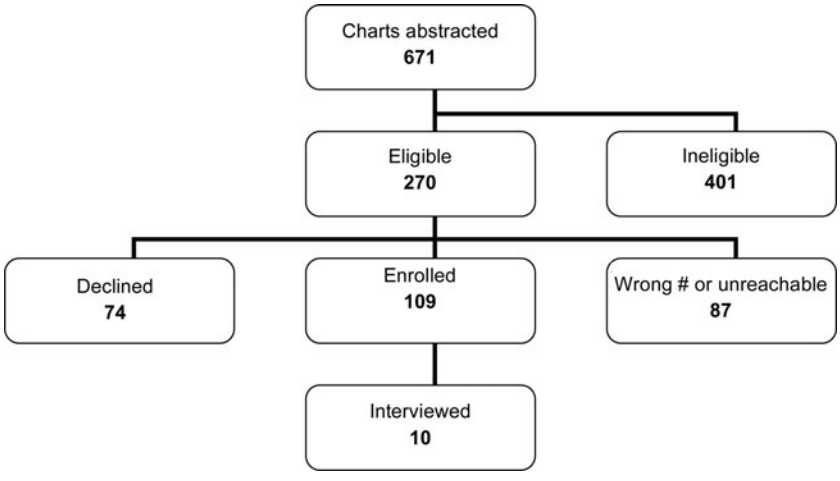


Figure 1. Study enrollment process: Practicality outcomes.

Table 1. Participant Characteristics (N = 267)

N	Enrolled (109)	Declined (74)	Unreachable (87)	Total (267)
Age, y (P =.06)				
51-60	65 (60%)	32 (43%)	50 (60%)	147 (55%)
61-70	44 (40%)	42 (57%)	34 (40%)	120 (45%)
Race (P = .53)				
White	32 (29%)	23 (31%)	27 (32%)	82 (31%)
Black	67 (62%)	46 (62%)	45 (54%)	158 (59%)
Other	10 (9%)	5 (7%)	12 (14%)	27 (10%)
Insurance status (P =.06)				
Public	72 (66%)	43 (58%)	39 (47%)	154 (58%)
Private	31 (28%)	29 (39%)	38 (45%)	98 (37%)
Uninsured	6 (6%)	2 (3%)	7 (8%)	15 (5%)
Last primary care visit (P = .0001)				
<1 y	97 (89%)	55 (74%)	53 (63%)	205 (77%)
>1 y	12 (11%)	19 (26%)	31 (37%)	62 (23%)
Body mass index (P = .52)				
Normal (<24.9)	17 (16%)	9 (12%)	11 (13%)	37 (14%)
Overweight (25-29.9)	27 (24%)	20 (27%)	21 (25%)	68 (25%)
Obese (≥30)	62 (57%)	42 (57%)	44 (52%)	148 (56%)
Not in chart	3 (3%)	3 (4%)	8 (10%)	14 (5%)
Current smoker[a]	29 (27%)	...	...	...
Depression[a]	31 (28%)	...	...	...

[a]Data are available only for enrolled subjects.

scheduled a primary care visit to follow up on their health priorities, and after 30 days, 54% of participants arrived for a visit with their primary care provider. Manual chart abstraction revealed that 64% of provider visits had documentation that 1 or more of participant's health priorities were addressed during that appointment.

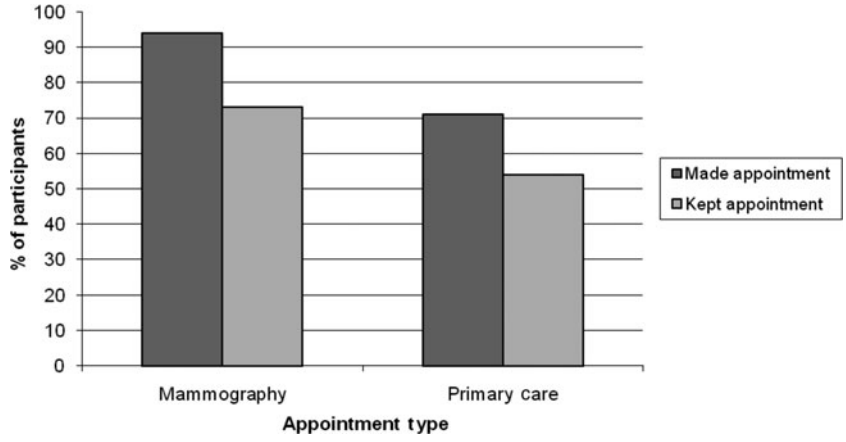


Figure 2. Made and kept appointments (N = 109).

Acceptability outcomes: The patient perspective

Overall, participants reported high satisfaction with the intervention. When asked what they most liked about the conversation with the navigator, participants most frequently noted the opportunity to speak with someone about their health and offering to schedule their appointments, especially concurrent appointments. Several also noted the value of sharing positive examples of their own self-care (eg, weight loss or quitting smoking). The majority of participants had nothing negative to say about the phone call, although one called it "excessive" and one said she would rather speak to someone face to face than over the phone.

Participants generally reported feeling very comfortable with and respected by the navigator (eg, "treated like a human being"). Several noted the importance of being heard by someone other than their doctor. One woman expressed it this way: "Navigators can help people who are in a shell—maybe they'll come out to her and speak to her about certain things. So I think it was a good idea." Another woman said that it would be good for some participants "to know that someone feels that they're special enough to come out and speak to them."

On average, all of the conversations about specific topics (smoking, depression, and weight) were rated above a helpfulness of 4.0, while those about smoking and receiving a mammogram were rated most favorably. All participants were favorable to the idea of combining the mammogram check-in with talk about other health issues. Some said that it gave them a chance to consider health-related issues that are important but that they had forgotten to pay attention to. Overall, participants reported liking the "holistic approach," saying that it was important to focus on the "mind and body" and "the whole picture."

The majority of participants reported that the priority list was helpful in preparation for the primary care visit. One participant stated, "I have a lot of different health issues, and it was very helpful for me to stop and think, well, which of those is the most important to me? You know, it just kind of helps me evaluate where I stand."

Only 2 participants accepted the navigator's offer to accompany her to the primary care visit. Another stated that introducing the navigator into the patient-provider relationship might risk conflicting the already-established rapport: "I don't want you to do that because [my provider has] been trying to get me to see a doctor. And I don't want her to think that I'm asking you and overstepping her."

Participants uniformly spoke positively about reminder calls provided by the navigator before both the mammography and primary care appointments. Several mentioned that the reminder was particularly useful because they tend to forget about scheduled appointments. One said, "I think the reminder was good because I don't like tests and stuff. Her reminding me—and she made the appointment, so that was the same day. So I got everything done. Otherwise, I probably would have said, forget it!"

During the follow-up call, the majority of respondents mentioned that the list of health priorities elicited by the patient herself was useful during the clinical visit with the primary care provider. In regard to the helpfulness of the priority list, 1 woman commented, "It helped me to stay on focus and to be able to do the things and tell the things that—you know because when you go to the doctor's office you go in there and you say, 'Yup, Yup,' [. . .]. Sometimes I do that and really don't say anything." Another remarked, "I think it helped me prioritize and pick the things that were most important in terms of talking to the doctor."

Acceptability outcomes: The navigator perspective

When compared with previous patient navigation approaches, the navigators felt that the incorporation of B-MI and chronic disease created a more personalized experience, which fostered open and honest conversations and a greater sense of patient empowerment. On a scale of 1 to 5, all of the navigators rated

their satisfaction with bundled patient navigation as 3 or higher. Navigators noted that the B-MI intervention gave them a better sense of what each patient needed, and provided them with the resources to address each patient's distinct health priorities. One navigator remarked, "When I reiterate or when I give them a summary or give them a feedback of what they've said, they seem to expand more. And I could feel them, like they take charge of it. A light bulb hits them."

Despite the numerous benefits of enhanced patient navigation, the navigators also identified challenges with the new approach. First, the navigators found that the intervention was unnecessarily excessive for some patients, in which case, the patients just wanted to schedule their appointments and did not want to engage in the B-MI intervention. They noted that the intervention proved to be less useful for the unanticipated high number of patients who had long-standing, positive relationships with their providers, and who felt confident voicing their own health concerns. The navigators uniformly found the depression screening to be an inappropriate screening tool to conduct over the phone with this patient population. They reported that it was often misunderstood or misinterpreted by patients, and these confusions and the stigma attached to depression resulted in some patients being less engaged and open after the depression screening took place.

The navigators also confronted systemic and logistic challenges. Above all, they felt that the commitment involved in this pilot project was too much to tack onto their day-to-day professional responsibilities already in place. They were frustrated by scheduling difficulties and a lack of patient/provider awareness of patient navigation. They recommend that this type of navigation be implemented within an environment that has a stronger foundation to support enhanced patient navigation.

Acceptability outcomes: The provider perspective

Generally speaking, the providers were satisfied with the enhanced patient navigation. Unfortunately, because of the size of their case load and the lapse of time between the patient visits and their interview, it was difficult for the providers to remember specific aspects of the patient navigation that were relevant to the patient visit. The providers were most positive about the navigators' role in eliciting the patient's health priorities. Three of the 4 providers interviewed agreed that many patients come to their visit and have a hard time prioritizing their health needs because they "tend to come in with a bunch of stuff." In the future, the providers recommended that patient navigation be used as a broad-based model, which can be tailored to address patients' specific psychosocial needs. The providers also felt that patient navigation could be a useful tool to support patients in complying with visits and routine health care maintenance.

DISCUSSION

The findings from this chronic disease navigation demonstration project targeting vulnerable participants confirm the feasibility of several aspects of a telephone-based community health worker intervention. The high prevalence of chronic disease in this population in need of a single-site cancer screening (ie, mammography) supports the need to adopt a more patient-centered approach like the one here. We found that almost all participants reached by telephone were willing to engage in the process of navigation and this translated into *effective* delivery of preventive health services for a population at risk for poor outcomes. However, the *practicality* of telephone outreach was limited because of language barriers and inability to reach patients despite multiple attempts. Patients were generally *acceptable* of telephone navigation; however, they did not readily accept face-to-face navigation activities. Navigators struggled with incorporating depression screening. These findings are critical to the planning process for care coordination programs that target similar populations.

Practicality

Perhaps most important is our finding that patient navigation—most frequently conducted in person and focused on a singular condition, such as breast cancer screening follow-up—can also be effectively offered via telephone and successfully combined with screening for carefully selected other health priorities and appointment making. In the context of conversations that were geared first to mammogram appointments, navigators were able to screen for overweight and smoking, both risk factors for breast cancer. Screening for depression proved more difficult, possibly related to the sensitivity of the topic and the awkwardness of the screening tool, particularly when administered by telephone. Navigators found it uncomfortable to ask what they perceived as very personal questions in this context. While some participants did not mind talking about depression, others appeared to be confused about what navigators were actually asking.

The practicality of this navigation program was limited, mainly on the basis of its reliance on telephone outreach. Navigators were able to reach only 68% of eligible patients by telephone, which is similar to telephone-based community health worker interventions targeting similar populations (Lasser et al., 2009; Myers et al., 2008; Percac-Lima et al., 2009; Weber & Reilly, 1997). Studies that were able to reach a larger percentage of patients frequently utilized additional recruitment strategies such as in-person recruitment and incentives for participation (Dietrich et al., 2006; Jandorf et al., 2005; Parker et al., 2007). We recommend that future studies consider additional recruitment strategies to compensate for the difficulties inherent in telephone-based outreach.

Efficacy

Once navigators engaged patients, using B-MI skills, the clinical outcomes were positive and comparable with other navigation studies. This study was conducted with a cohort of women who had not received a mammogram in 18 to 24 months and were thus nonadherent with screening. The ability of the intervention to result in 73% of nonadherent participants completing mammograms is impressive when compared with other intervention studies targeting populations overdue for screening. In a study by Weber and Reilly (1997), 41% of participants overdue for mammograms completed a mammogram in the 10-month intervention period. A study by Burhansstipanov et al. achieved a 55% mammogram completion rate among women who had not been screened within 18 months (Burhansstipanov et al., 2010), and in a study by Fernandez et al., 41% of women who had not had a mammogram in 12 months completed a screening within the 6-month follow-up period (Fernandez et al., 2009). However, because of the lack of a direct comparison group, generalizability of efficacy data is limited.

The full effect of the intervention may also have been limited by the failure of patient health priorities reaching the intended target, the primary care provider. The intervention was intended to prepare participants for their upcoming primary care visit, by discussing priority health concerns in advance. Navigators elicited from patients their list of priority health concerns, a list that was to serve as a springboard for conversation with their provider. While participants reported benefits from generating the list, it did not, in the end, arrive at the providers' offices/records as planned or (be perceived to) exert an impact on the visit. It is likely that this failure was related to the low level of awareness among providers about the intervention, and the long-time lag between the navigation conversation and the primary care visit, and between the primary care visit and the follow-up interview.

Acceptability

Once enrolled, patients talked freely about their health worries and strategies for self-care and reported high satisfaction with the conversation. The fact that they did not readily accept face-to-face navigation services when offered was interesting and may well be related to the finding that the relationships between primary care providers and patients in

this study tend to be long-standing, strong, and open. Thus, the presence of a navigator and even the added information from a patient navigator may not be as critical as it might be in settings where such continuity and strength in provider-patient relationships are not the norm. In fact, providers who participated in follow-up interviews indicated little knowledge that the patient navigation and screening had or had not occurred. When asked about it hypothetically, they were most likely to comment positively on the value of a preset patient-generated priority list, particularly for patients with complex psychosocial needs.

Key lessons learned and recommendations for future telephone-based navigation in this setting include the following:

- Find ways to engage and inform providers early on and in more depth about the intervention and its intent to better prepare patients and providers for their primary care visit.
- Expand hours in which the intervention can be conducted to increase the likelihood of reaching eligible individuals with fewer phone calls.
- Send the patient's priority list to the provider in a clear and recognized way (through the patient, through the navigator, or through the EMR). The priority list also could be effectively revisited when the navigator calls back to schedule or reschedule visits.
- Exclude depression screening in this context.
- Make explicit the "same day mammography/primary care visit" goal for each patient.

LIMITATIONS

Because of the focus on quality improvement and feasibility, we did not include a direct comparison group in this demonstration project. While not ideal, this design is consistent with other feasibility studies that aim to determine whether an intervention should be recommended for efficacy testing (Bowen et al., 2009) and the efficacy demonstrated here is consistent with other navigation evaluation literature. This study was also limited by the restricted inclusion criteria to English speakers only, which is not fully representative of the diverse patient population served at this safety net institution. This limitation underscores the reality of providing culturally appropriate care in resource poor settings and calls for future planning that prioritizes multilingual health workers.

CONCLUSION

As patient navigation becomes a more prevalent form of community health work, in the era of patient-centered health care reform it is crucial to expand navigation beyond the traditional disease-specific cancer setting and provide navigators with tools to effectively address multiple chronic conditions. This study supports the feasibility of chronic disease patient navigation in primary care with some key lessons for program planning and future efficacy testing. This type of feasibility data is critical to incorporating community health workers into the emerging patient-centered medical home model to ensure the delivery of quality care to the most vulnerable populations.

REFERENCES

American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, American Osteopathic Association. (2007). *Joint principles of the patient-centered medical home*. Retrieved December 22, 2010, from http://www.aafp.org/pcmh/principles.pdf

American College of Surgeons. (2009). *Commission on cancer. Cancer program standards 2009*. Retrieved January 25, 2011, from http://www.facs.org/cancer/coc/cocprogramstandards.pdf

Apodaca, T. R., & Longabaugh, R. (2009). Mechanisms of change in motivational interviewing: A review and

preliminary evaluation of the evidence. *Addiction*, *104*(5), 705–715.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychology Review*, *84*, 191–215.

Bastani, R., Mojica, C. M., Berman, B. A., & Ganz, P. A. (2010). Low-income women with abnormal breast findings: Results of a randomized trial to increase rates of diagnostic resolution. *Cancer Epidemiology Biomarkers Prevention*, *19*(8), 1927–1936.

Battaglia, T. A., Roloff, K., Posner, M. A., & Freund, K. M. (2007). Improving follow-up to abnormal breast cancer screening in an urban population. A patient navigation intervention. *Cancer*, *109*(2, Suppl.), 359–367.

Bernstein, E., Bernstein, J., Feldman, J., Fernandez, W., Hagan, M., Mitchell, P., . . . Owens, P. (2007). An evidence based alcohol screening, brief intervention and referral to treatment (SBIRT) curriculum for emergency department (ED) providers improves skills and utilization. *Substance Abuse*, *28*(4), 79–92.

Bernstein, J., Bernstein, E., Tassiopoulos, K., Heeren, T., Levenson, S., & Hingson, R. (2005). Brief motivational intervention at a clinic visit reduces cocaine and heroin use. *Drug and Alcohol Dependence*, *77*(1), 49–59. Retrieved January 24, 2011, from The BNI-ART Institute website: http://www.ed.bmc.org/sbirt/

Bernstein, E., Edwards, E., Dorfman, D., Heeren, T., Bliss, C., & Bernstein, J. (2009a). Screening and brief intervention to reduce marijuana use among youth and young adults in a pediatric emergency department. *Academic Emergency Medicine*, *16*(11), 1174–1185.

Bernstein, E., Topp, D., Shaw, E., Girard, C., Pressman, K., Woolcock, E., & Bernstein, J. (2009b). A preliminary report of knowledge translation: Lessons from taking screening and brief intervention techniques from the research setting into regional systems of care. *Academic Emergency Medicine*, *16*(11), 1225–1233.

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... Fernandez, M. (2009). (2009). How we design feasibility studies. *American Journal of Preventive Medicine*, *36*(5), 452–457.

Brug, J., Spikmans, F., Aartsen, C., Breedveld, B., Bes, R., & Fereira, I. (2007). Training dietitians in basic motivational interviewing skills results in changes in their counseling style and in lower saturated fat intakes in their patients. *Journal of Nutrition Education Behavior*, *39*(1), 8–12.

Burhansstipanov, L., Dignan, M. B., Schumacher, A., Krebs, L. U., Alfonsi, G., & Apodaca, C. C. (2010). Breast screening navigator programs within three settings that assist underserved women. *Journal of Cancer Education*, *25*(2), 247–252.

Chen, L. A., Santos, S., Jandorf, L., Christie, J., Castillo, A., Winkel, G., & Itzkowitz, S. (2008). A program to enhance completion of screening colonoscopy among urban minorities. *Clinical Gastroenterology Hepatologie*, *6*(4), 443–450.

Dietrich, A. J., Tobin, J. N., Cassells, A., Robinson, C. M., Greene, M. A., Sox, C. H., . . . Younge, R. G. (2006). Telephone care management to improve cancer screening among low-income women: A randomized, controlled trial. *Annals of Internal Medicine*, *144*(8), 563–571.

Dohan, D., & Schrag, D. (2005). Using navigators to improve care of underserved patients: Current practices and approaches. *Cancer*, *104*(4), 848–855.

Ell, K., Vourlekis, B., Lee, P. J., & Xie, B. (2007). Patient navigation and case management following an abnormal mammogram: A randomized clinical trial. *Preventive Medicine*, *44*(1), 26–33.

Fernández, M. E., Gonzales, A., Tortolero-Luna, G., Williams, J., Saavedra-Embesi, M., Chan, W., & Vernon, S. W. (2009). Effectiveness of Cultivando la Salud: A breast and cervical cancer screening promotion program for low-income Hispanic women. *American Journal of Public Health*, *99*(5), 936–943.

Ferrante, J. M., Cohen, D. J., & Crosson, J. C. (2010). Translating the patient navigator approach to meet the needs of primary care. *Journal of American Board Family Medicine*, *23*(6), 736–744.

Ford, M. E., Havstad, S., Vernon, S. W., Davis, S. D., Kroll, D., Lamerato, L., & Swanson, G. M. (2006). Enhancing adherence among older African American men enrolled in a longitudinal cancer screening trial. *Gerontologist*, *46*(4), 545–550.

Freeman, H. P., Muth, B. J., & Kerner, J. F. (1995). Expanding access to cancer screening and clinical follow-up among the medically underserved. *Cancer Practice*, *3*(1), 19–30.

Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion*, *11*(2), 87–98.

Greenfield, S., Kaplan, S., & Ware, J. E., Jr. (1985). Expanding patient involvement in care: Effects on patient outcomes. *Annals of Internal Medicine*, *102*(4), 520–528.

Greenfield, S., Kaplan, S. H., Ware, J. E., Jr., Yano, E. M., & Frank, H. J. (1988). Patient participation in medical care: Effects on blood sugar control and quality of life in diabetes. *Journal of General Internal Medicine*, *3*(5), 448–457.

Guadagnolo, B. A., Boylan, A., Sargent, M., Koop, D., Brunette, D., Kanekar, S., ..., Petereit, D. G. (2010). Patient navigation for American Indians undergoing cancer treatment: Utilization and impact on care delivery in a regional healthcare center. *Cancer*. doi:10.1002/cncr.25823

Hardeman, W., Johnston, M., Johnston, D. W., Bonetti, D., Warham, N. J., & Kinmonth, A. L. (2002). Application of the theory of planned behaviour in behaviour change interventions: A systematic review. *Psychology and Health*, *17*(2):123–158.

Health Resources and Services Administration, U.S. Department of Health and Human Services. (2007).

Community health worker national workforce study. Retrieved January 17, 2011, from http://bhpr.hrsa.gov/healthworkforce/chw/

Hettema, J., Steele, J., & Miller, W. R. (2005). Motivational interviewing. *Annual Review of Clinical Psychology*, *1*, 91–111.

Jandorf, L., Gutierrez, Y., Lopez, J., Christie, J., & Itzkowitz, S. H. (2005). Use of a patient navigator to increase colorectal cancer screening in an urban neighborhood health clinic. *Journal of Urban Health*, *82*(2), 216–224.

Kaner, E. F., Beyer, F., Dickinson, H. O., Pienaar, E., Campbell, F., Schlesinger, C., . . . Burnand,B. (2007). Effectiveness of brief alcohol interventions in primary care populations. *Cochrane Database System Review*, (2), CD004148.

Kaplan, S. H., Gandek, B., Greenfield, S., Rogers, W., & Ware, J. E. (1995, December). Patient and visit characteristics related to physicians' participatory decision-making style. Results from the Medical Outcomes Study. *Medical Care*, *33*(12), 1176–1187.

Kaplan, S. H., Greenfield, S., & Ware, J. E., Jr. (1989, March). Assessing the effects of physician–patient interactions on the outcomes of chronic disease. *Medical Care, 27*(3, Suppl.), S110–S127.

Kressin, N. R., Nunn, M. E., Singh, H., Orner, M. B., Pbert, L., Hayes, C., . . . Henshaw, M. M. (2009). Pediatric clinicians can help reduce rates of early childhood caries: Effects of a practice based intervention. *Medical Care*, *47*(11), 1121–1128.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire–2: Validity of a two-item depression screener. *Medical Care*, *41*(11), 1284–1292.

Lasser, K. E., Murillo, J., Medlin, E., Lisboa, S., Valley-Shah, L., Fletcher, R. H., & Ayanian, J. Z. (2009). A multilevel intervention to promote colorectal cancer screening among community health center patients: Results of a pilot study. *BMC Family Practice*, *10*, 37.

Lowe, B., Kroenke, K., & Grafe, K. (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, *58*(2), 163–171.

Lubetkin, E. I., Lu, W. H., Krebs, P., Yeung, H., & Ostroff, J. S. (2010). Exploring primary care providers' interest in using patient navigators to assist in the delivery of tobacco cessation treatment to low income, ethnic/racial minority patients. *Journal of Community Health*, *35*(6), 618–624.

Marvel, M. K., Epstein, R. M., Flowers, K., & Beckman, H. B. (1999, January 20). Soliciting the patient's agenda: Have we improved? *Journal of the American Medical Association*, *281*(3), 283–287.

Mason, M. J. (2009). Rogers redux: Relevance and outcomes of motivational interviewing across behavioral problems. *Journal of Counseling and Development*, *87*, 357–362.

Myers, R. E., Hyslop, T., Sifri, R., Bittner-Fagan, H., Katurakes, N. C., Cocroft, J., . . . Wolf, T. (2008). Tailored navigation in colorectal cancer screening. *Medical Care*, *46*(9, Suppl. 1), S123–S131.

Ockene, I. S., Hebert, J. R., Ockene, J. K., Merriam, P. A., Hurley, T. G., & Saperia, G. M. (1996). Effect of training and a structured office practice on physician-delivered nutrition counseling: The Worcester-Area Trial for Counseling in Hyperlipidemia (WATCH). *American Journal of Preventive Medicine*, *12*(4), 252–258.

Ockene, I. S., Hebert, J. R., Ockene, J. K., Saperia, G. M., Stanek, E., Nicolosi, R., . . . Hurley, T. G. (1999). Effect of physician-delivered nutrition counseling training and an office-support program on saturated fat intake, weight, and serum lipid measurements in a hyperlipidemic population: Worcester Area Trial for Counseling in Hyperlipidemia (WATCH). *Archives of Internal Medicine*, *159*(7), 725–731.

Parker, D. R., Windsor, R. A., Roberts, M. B., Hecht, J., Hardy, N. V., Strolla, L. O., & Lasater, T. M. (2007). Feasibility, cost, and cost-effectiveness of a telephone-based motivational intervention for underserved pregnant smokers. *Nicotine and Tobacco Research*, *9*(10), 1043–1051.

Percac-Lima, S., Grant, R. W., Green, A. R., Ashburner, J. M., Gamba, G., Oo, S., . . . Atlas, S. J. (2009). A culturally tailored navigator program for colorectal cancer screening in a community health center: A randomized, controlled trial. *Journal of General Internal Medicine*, *24*(2), 211–217.

Romano, J. L., & Netland, J. D. (2008). The application of the theory of reasoned action and planned behavior to prevention science in counseling psychology. *Counseling Psychologist*, *36*(5), 777–806.

Scholle, S. H., Saunders, R. C., Tirodkar, M. A., Torda, P., & Pawlson, L. G. (2010). Patient-centered medical homes in the United States. *Journal of Ambulatory Care Management*, *34*(1), 20–32.

Stange, K. C., Nutting, P. A., Miller, W. L., Jaen, C. R., Crabtree, B. F., Flocke, S. A., & Gill, J. M. (2010). Defining and measuring the patient-centered medical home. *Journal of General Internal Medicine*, *25*(6), 601–612.

Stein, M. D., Herman, D. S., & Anderson, B. J. (2009). A motivational intervention trial to reduce cocaine use. *Journal of Substance Abuse Treatment*, *36*(1), 118–125.

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *132*(2), 249–268.

Weber, B. E., & Reilly, B. M. (1997). Enhancing mammography use in the inner city. A randomized trial of intensive case management. *Archives of Internal Medicine*, *157*(20), 2345–2349.

Whitlock, E. P., Polen, M. R., Green, C. A., Orleans, T., & Klein, J. (2004). Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: A summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*, *140*(7), 557–568.

Connecting Boston's Public Housing Developments to Community Health Centers: Who's Ready for Change?

Tracey A. Battaglia, MD, MPH[1], Samantha S. Murrell, MPH[1], Sarah Gees Bhosrekar, MPH[2], Sarah E. Caron, MPH[1], Deborah J. Bowen, PhD[2], Eugenia Smith[3], Richard Kalish, MD[4], Jo-Anna Rorie, CNM, MSN, MPH[2]

(1) Women's Health Unit, Section of General Internal Medicine, Department of Medicine, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine; (2) Department of Community Health Science, Boston University School of Public Health; (3) Community Committee for Health Promotion; (4) Boston University Clinical and Translational Sciences Institute

Submitted 1 August 2011, revised 12 February 2012, Accepted 2 March 2012. This work was supported by Cooperative Agreement Number 5U48 DP001922-02 from the Centers for Disease Control and Prevention and also supported by the Boston University-Clinical Translational Sciences Institute BU-BRIDGE (NIH# UL1RR025771).

Abstract

Background: Despite close proximity to community health centers, public housing residents are at increased risk of uncontrolled chronic disease, in part because of underutilization of routine health care.

Objectives: To assist in program planning, the Partners in Health and Housing Prevention Research Center (PHH-PRC) used the Community Readiness Model to compare readiness of public housing developments and community health centers to address community-identified health priorities. The model assumes that program success to affect change depends on matching the community's level of readiness to address the issue.

Methods: Key respondent interviews were conducted across 15 communities: Eight housing developments and seven health centers. Interviews were scored across six dimensions on an anchored, 9-point scale and averaged to provide a composite readiness score. Higher scores indicate increasing levels of readiness. Interview transcripts were reviewed for consistent themes.

Results: Health centers scored significantly higher (mean, 5.88) than housing developments (mean, 3.33), corresponding with the Preparation stage of readiness compared with the Vague Awareness stage, respectively. Both scored highest in Existing Programs and Resources and lowest in Knowledge of Efforts. Qualitative analysis revealed a lack of existing partnerships between housing developments and health centers as well as significant social barriers preventing housing residents from engaging in care.

Conclusion: We found a mismatch in readiness to address community health priorities. Although health centers have programs to address health issues, community awareness of programs is limited and barriers to engaging in care persist. The model provided a useful tool for engaging communities into shared program planning.

Keywords

Community-based participatory research, community health partnerships, urban health, urban health services, health disparities, health promotion, health outcomes

Most of the 2 million Americans living in public housing fall below the federal poverty level and face significant social, economic, and physical stressors that negatively impact their health.[1] This is demonstrated by the disproportionately high burden of chronic health problems among public housing residents nationwide, including high rates of asthma, cancer, oral health issues, obesity, hypertension, and diabetes. As a result, public housing residents are more likely than the general population to suffer early morbidity and mortality.[1-5] In Boston, Massachusetts, nearly 10% of residents live in public housing or receive rental assistance.[6] Research has shown that these public housing residents are over three times more likely to report fair or poor health status compared with other Boston residents;

they are more likely to have ever been diagnosed with high cholesterol, hypertension, and diabetes; and report higher current rates of asthma, obesity, smoking, depressive tendencies, disability, and insufficient physical activity.[7] These ongoing health disparities support the need for innovative interventions targeting this vulnerable population.

Public housing residents live in closed multigenerational communities in close physical and social proximity. This creates opportunity for proactive community-based interventions that aim to address problems in the context of community strengths while respecting cultural diversity.[8-10] Engagement of the community is critical to translating scientific advances in healthcare to improved health status, yet numerous barriers exist for vulnerable populations to participate in scientific inquiry, including lack of access and opportunity.[11-13] These challenges limit our understanding of the causes of health disparities and of effective interventions to reduce these disparities among Boston's public housing residents. A community-based participatory research approach is one way to ensure participation of this population in intervention planning.

The PHH-PRC was established in Boston in 2001 as 1 of 33 Prevention Research Centers nationwide, funded by the U.S. Centers for Disease Control and Prevention to create community–academic partnerships aimed at reducing the burden of disease and promoting health among at-risk populations. The PHH-PRC is the only such collaboration that focuses exclusively on public housing, engaging residents in activities and research designed to improve their health and well-being and thus reduce health disparities.[14] The collaboration includes four equal partners: Boston University School of Public Health, Boston Public Health Commission, Boston Housing Authority, and the Community Committee for Health Promotion, representing the residents of Boston Housing Authority's family housing developments. Prior work by the PHH-PRC found that 65% of residents were found to have evidence of at least one life-threatening chronic illness.[15] When followed for 3 months after their screening, only 15% of those with chronic disease had evidence of a follow-up medical appointment,[15] despite close proximity to numerous federally qualified health centers and academic medical centers in Boston. This finding calls for collaborative approaches between the community and local clinical care sites.

Table 1. Community Readiness Model Stages[18-21]

Stage	Description	Goal
1. No awareness	Issue is not generally recognized by the community or leaders as a problem.	Raise awareness of the issue.
2. Denial/resistance	At least some community members recognize that it is a concern, but there is little recognition that it might be occurring locally.	Raise awareness that the issue exists in this community.
3. Vague awareness	Most feel that there is a local concern, but there is no immediate motivation to do anything about it.	Raise awareness that the community can do something.
4. Preplanning	There is clear recognition that something must be done, and there may even be a group addressing it. However, efforts are not focused or detailed.	Raise awareness with concrete ideas to combat condition.
5. Preparation	Active leaders begin planning in earnest. Community offers modest support of efforts.	Gather existing information with which to plan strategies.
6. Initiation	Enough information is available to justify efforts. Activities are underway.	Provide community-specific information.
7. Stabilization	Activities are supported by administrators or community decision makers. Staff are trained and experienced.	Stabilize efforts and programs.
8. Confirmation/expansion	Efforts are in place. Community members feel comfortable using services, and they support expansions. Local data are regularly obtained.	Expand and enhance services.
9. High level of community ownership	Detailed and sophisticated knowledge exists about prevalence, causes, and consequences. Effective evaluation guides new directions. Model is applied to other issues.	Maintain momentum and continue growth.

As preliminary research for interventions that link public housing residents with primary care services, we turned to the Community Readiness Model to assess the level of readiness of the "community" to address a community-identified health priority. The model, developed by the Tri-Ethnic Center for Prevention Research at Colorado State University, is a widely used model for community change[16] that integrates a community's culture, resources, and level of readiness to guide the complex process of community change. The underlying assumption of the model is that overall success of programs to affect change depends on matching the program with a community's level of readiness to address the chosen issue.[17] Community readiness theory combines principles of individual psychological readiness for change (trans-theoretical model) with the principles of community development to provide a framework for implementing changes in community health services,[18,19] then offers practical tools to design, implement, and evaluate programs in community settings.[17] The model defines nine stages of community readiness ranging from no awareness of the problem to a high level of community ownership and action (Table 1). Each stage is characterized by a set of criteria and accompanied by a set of specific goals that facilitate advancing the level or stage of readiness. Originally designed to target substance abuse and drug prevention by the Tri-Ethnic Center,[17,19] the model has since emerged as a tool in other areas,[20] including obesity,[21,22] HIV,[23] and clinical trial enrollment.[24]

The aim of this study was to understand the level of readiness for each community partner (housing developments *and* health centers) to engage in health promotion planning, as a preliminary step toward successful implementation of programs that would engage residents into clinical care. We target both public housing developments and community health centers as "communities" of interest. The aims of this research were to (1) compare the readiness of housing developments versus community health centers to address a chosen health priority, and (2) identify community-specific barriers relevant to bringing residents of public housing developments into primary care services within the community health centers.

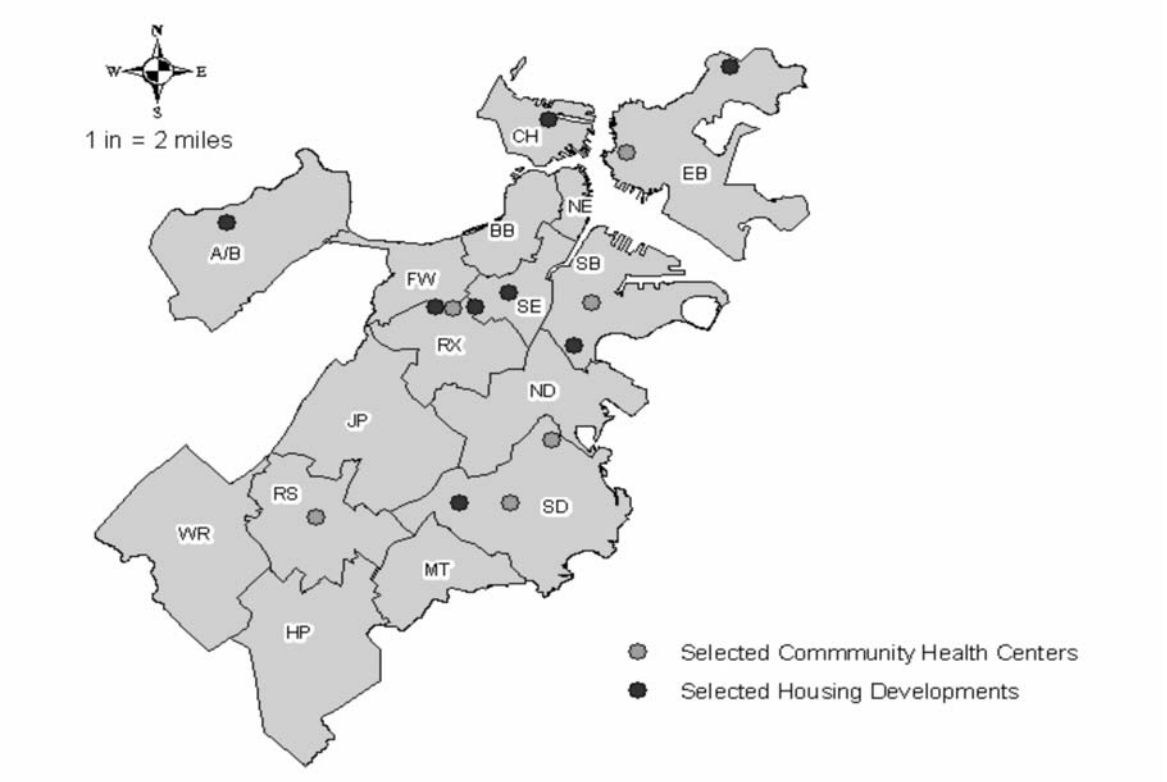


Figure 1. Location of Study Sites by Boston Neighborhoods

Boston neighborhood names: A/B = Allston-Brighton; BB = Back Bay; CH = Charlestown; EB = East Boston; FW = Fenway; HP = Hyde Park; JP = Jamaica Plain; MT = Mattapan; ND = North Dorchester; RS = Roslindale; RX = Roxbury; SB = South Boston; SD = South Dorchester; SE = South End; WR = Roxbury.

METHODS

Utilizing principles of the community readiness theory, we conducted readiness assessments across 15 different communities, including eight public housing developments in Boston and seven federally qualified community health centers in close proximity to the chosen housing developments (Figure 1). Key respondent interviews at each community site were conducted between May and October 2010. The Boston University Medical Center Institutional Review Board approved this study.

Study Sites

For this project, we define a "community" as a group that shares a common residence or clinical care site, as defined by a discrete public housing development or community health center. Rather than a geographic area or cultural group, we view a public housing development or community health center as a distinct community with its own identity.[25] We included health centers and housing developments within the same geographic area as discrete communities because they have independent governing structures, health priorities, and ways of addressing those priorities.

There are a total of 26 family public housing developments throughout Boston that provide quality, public, and affordable housing for low- and moderate-income individuals and families from diverse racial/ethnic backgrounds. In 2009, 11% of Boston Public Housing Residents were White, 32% were Black, 47% were Hispanic, and 10% "other" racial/ethnic groups. The average household income for these residents is less than $14,000 per year.[26] Limited resources required the PHH-PRC Steering Committee to identify eight family housing developments to participate. Criteria for selection included high burden of chronic disease, size of development, racial/ethnic variability, and geographic variability.

The Clinical and Translational Sciences Institute at Boston University, BU BRIDGE, provided financial support for inclusion of seven of their 15 primary-affiliate federally qualified health centers in the assessment. These community-based, outpatient health, and multi-service centers located throughout Boston's neighborhoods provide comprehensive care to Boston's most vulnerable populations. In 2007, these centers provided care to nearly 200,000 individual patients (23% White, 35% Black, 19% Hispanic, 4% Asian/Pacific Islander) and nearly 80% living below 200% of the Federal Poverty Level.[27] We chose health centers in close proximity to housing developments to compare readiness of institutions working within the same geographic area, and thus serving similar populations. The main criterion for selection of the 7 health centers was geographic proximity to the housing developments (Figure 1). Together, the eight housing developments and seven health centers comprised the 15 communities included in this assessment.

Health Priority

Because readiness assessment is issue-specific and we wanted to ensure community engagement in the research process, the research team asked leadership from each of the 15 communities to choose one priority health topic for the assessment. Housing developments were represented collectively by the PHH-PRC partners who utilized data from their annual resident survey to identify multiple chronic diseases of concern, including obesity, diabetes, high cholesterol, hypertension, and smoking. PHH-PRC partners agreed to combine these chronic diseases into the broad category of cardiovascular disease or "heart health" for their assessments. Health center leadership, which included medical and administrative staff from each of the seven participating sites, was asked to identify the health priority for their individual health center. Priorities differed across health centers, though most also related to "heart health"; three chose obesity, one childhood obesity, one hypertension, one asthma, and one substance abuse.

Study Subjects

Readiness methodology requires four to six key respondent interviews per community. Key respondents were identified by leadership from their community, The Community Committee for Health Promotion for housing developments and medical directors or executive directors of the health center. Criteria for key respondents included adult community members knowledgeable about ongoing efforts around the chosen health priority and diversity across racial/ethnic groups representative of the larger community.

Interview Guide

A standardized, open-ended interview guide was adapted to measure the six dimensions of readiness: (1) Community efforts: To what extent are there efforts, programs, and policies

that address the issue? (2) Community knowledge of efforts: To what extent do community members know about local efforts and their effectiveness? Are the efforts accessible to all segments of the community? (3) Leadership: To what extent are appointed leaders and influential community members supportive of the issue? (4) Community climate: What is the prevailing attitude in the community toward the issue? (5) Community knowledge of the issue: To what extent do community members know about the causes of the problem, consequences, and how it impacts the community? and (6) Resources related to the issue: To what extent are local resources (people, time, money) available to support the efforts? PHH-PRC partners contributed to the modification of an existing interview guide to target the issues identified by each community, addressed language and literacy concerns, and added additional questions that targeted partnerships between housing developments and health centers. Self-reported socio-demographic variables including gender, age, race/ethnicity, and role in community were included along with 40 open-ended readiness questions. Interview guides were piloted in March 2010 with two housing development residents and one health center staff and modified accordingly. (For a copy of the interview guide see the on-line appendix http://muse.jhu.edu/journals/progress_in_community_health_partnerships_research_educaton_and_action/v006/6.3.battaglia02_supp01.pdf)

Data Collection

After obtaining written informed consent, four to six in-person interviews were conducted at each of the 15 communities by two authors (SM, SGB), each of whom completed in-person training by Tri-Ethnic Center staff. Each one-on-one interview lasted between 45 and 75 minutes, was audio-taped, and sent via a secure, password-protected website to the Tri Ethnic Center for transcription and primary analysis.

Community Readiness Score

Using a previously validated scoring system,[17,18] experienced Tri-Ethnic Center staff scored each community across six the dimensions of readiness on an anchored 9-point scale, with higher scores indicating higher levels of readiness. Interviews were reviewed and scored independently by two reviewers, who then met to discuss scores that did not match to reach consensus. Each question ties to certain dimensions, scores are given per dimension, not per question. Scores for each dimension are averaged to give each community an overall readiness score,[16] and by convention are rounded down to determine the stage of readiness. As shown in Table 1, scores range from 1 to 9 corresponding to each stage: (1) No Awareness, (2) Denial/Resistance, (3) Vague Awareness, (4) Preplanning, (5) Preparation, (6) Initiation, (7) Stabilization, (8) Confirmation/Expansion, and (9) High Level of Community Ownership.

Analyses

We used descriptive statistics to provide a summary of participant characteristics using SAS version 9.1 (SAS, Inc., Cary, NC). A community readiness score was calculated for each community and an independent *t* test was used to compare scores between communities. Three members of the research team (SM, TB, SGB) conducted a thematic analysis of the transcribed interviews to systematically organize observations in the data that were not represented in the readiness scale.[28] Investigators independently reviewed transcripts to identify reoccurring themes, and then discussed independent findings in a group forum. Reoccurring themes were then grouped into agreed upon broader thematic categories. As demonstrated by previous research utilizing this model, this additional qualitative analysis serves to enhance quantitative results by fostering a more thorough understanding of the unique factors that contribute to a community's level of readiness.[29]

RESULTS

We interviewed a total of 78 key respondents (43 housing development residents and 36 health center staff). More than half (73%) were female (Table 2). Key respondents from the health centers were majority White (67%), 17% Black, and 14% Asian/Pacific Islander. Conversely, housing development key respondents were largely (86%) racial/ethnic minorities: 58% Black, 28% Hispanic, and 9% American Indian/Alaska Native. Overall, health center key respondents were younger than housing development residents; however, all adult age groups were well-represented in both community settings. By study design, key respondents included management, leadership, lay, and outside members affiliated with the community.

Community Readiness Scores

Overall, community readiness scores were significantly higher for health centers than housing developments (mean,

5.8 vs. 3.33, respectively; $p < .0001$). Table 3 also demonstrates that there was little variability in scores across the 2 types of community settings, such that health center scores ranged from 4.98 to 7.48 (SD 0.91), while housing development scores ranged from 2.58 to 4.24 (SD 0.65). A score of 5 at the health centers indicates the Preparation Stage, meaning that active community leaders have begun planning in earnest, but the community itself (the health center) offers modest support of these efforts. According to the Community Readiness Model guide, to advance health promotion around the priority issue, the goal for these communities is to continue to gather pertinent information and work with key leaders and influential community members to introduce information about the issue and ongoing efforts to the target population (Table 1).[16] Examples of stage-specific strategies for health centers include conducting patient interviews or surveys to identify service gaps, improving existing services to address those gaps, and identifying key places to post information. For housing developments, an overall score of 3 indicates the Vague Awareness Stage, meaning that there is local concern but no immediate motivation to take action around heart health. The goal for efforts in these communities is to raise awareness that the community can do something about this issue. One recommended strategy for advancing heart health promotion for these communities is educating existing small groups within the developments about heart health.

Figure 2 depicts the mean readiness scores by dimension, comparing means for health centers with housing development communities. Despite the overall difference in stage of readiness between these two types of community settings, both health centers and housing developments scored highest in existing programs and policies to address their health priority (mean of 6.65 vs. 4.01) and in the resources available to address the issue (mean of 5.98 vs. 3.97). Health centers scored lowest in leadership (mean, 5.43) and community knowledge of efforts (mean, 5.57), whereas housing developments scored lowest in community knowledge of efforts and community knowledge of the issue (with means of 2.58 and 2.76, respectively).

Table 2. Key Respondent Demographics ($N = 79$)

	CHCs ($n = 36$)		Housing ($n = 43$)	
	n	(%)	*n*	(%)
Gender				
Male	10	(28)	10	(23)
Female	26	(72)	32	(74)
Unspecified	0	(0)	1	(2)
Race*				
White	24	(67)	8	(19)
Hispanic/Latino	0	(0)	12	(28)
African American/Black	6	(17)	25	(58)
American Indian Alaska Native	1	(3)	4	(9)
Asian/Pacific Islander	5	(14)	0	(0)
Age (yrs)				
19–34	7	(20)	10	(23)
35–54	22	(61)	15	(35)
≥55	7	(19)	18	(42)
Role in the community				
Management	5	(14)	8	(19)
Clinical/resident leader	9	(25)	15	(35)
Provider/resident	10	(28)	18	(42)
Support staff/local nonprofit	12	(33)	2	(5)

* Percentages do not equal 100 because some participants selected multiple race/ethnicities.

Qualitative Findings

Several themes pertaining to programmatic planning were identified. In developments, two key themes emerged: (1) Barriers to heart health and (2) variability in site of care. In health centers, the major theme identified was a lack of targeted outreach efforts to housing developments.

Barriers to Heart Health. Housing development key respondents repeatedly report barriers to health that are attributable to basic life needs, including violence/safety, substance abuse, financial issues, insurance, and housing conditions. The following excerpts from two key respondents at different housing developments demonstrate this point:

Table 3. Summary of Overall Community Readiness Scores

Community	Mean Score	Range	Standard Deviation
Health Centers	5.88	4.98–7.48	0.91
Housing Developments	3.33	2.58–4.24	0.65

> The hallways are not clean. There's spit. There's vermin. There's all type of bad, disgusting things in there. . . . [Heart health] would be people's main concern because they want to be healthy, but come out and see that spit in front of your door; you might not think about jogging right after.
>
> It's a concern, but it's not at the front of people's minds. Gun violence is a big concern . . . you can see violence.

Site of Care. The second theme that emerged from the housing development key respondents was the tremendous variability noted in where residents reported receiving their care. Across the 43 housing development respondents, a total of 24 different sites of care were noted, and included both local community health centers and several academic medical centers. At one housing development, 13 different clinical sites were referenced as sites of care for community members.

Outreach to Developments. The final theme that emerged came from review of transcripts from health center key respondents. None of these key respondents felt that their current programming specifically targeted or reached out to their local housing development residents. While they did express an interest in working with these communities, they were largely unaware of any current initiatives, as these two excerpts demonstrate:

> We don't have any specific relationships [with local housing developments]. We just serve people that live in local housing projects, but we do not have any coordinated efforts that I'm aware of.
>
> We have patients who come in from . . . the housing projects, but there isn't a partnership.

DISCUSSION

This is the first study to use community readiness as a means to galvanize two unique community settings into action to address common health priorities among at-risk, underserved, urban populations. Our findings confirm a mismatch in health priorities as well as the level of readiness to address those health priorities between housing developments and community health centers. Although the community health centers report existing programs to address their priority health issues, community awareness of programs is limited in both settings. Further, housing developments demonstrate a local concern, but there is no immediate motivation to do anything about it, in part owing to perceived barriers to care. We found the readiness assessment to be a useful tool for engaging these communities in conversation and shared program planning by providing specific action-oriented guidance on stepwise program planning.

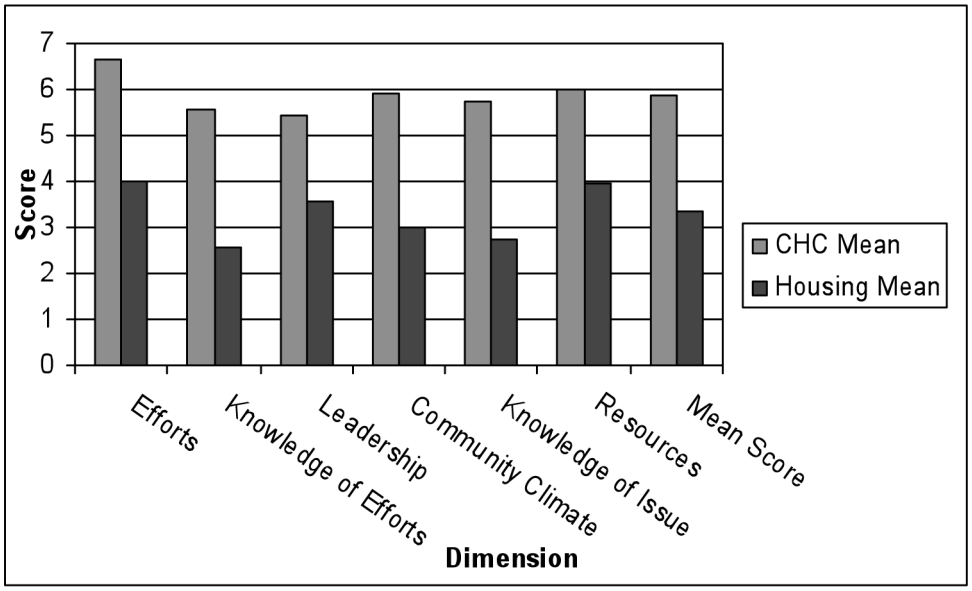


Figure 2. Mean Readiness Scores to Address Community-Identified Health Priorities by Dimension: Health Centers Versus Housing Developments

Each community was given the opportunity to identify its own health priority. Although there was much overlap in identified priorities, there was some disconnect between the two settings as reflected by one health center choosing asthma and another substance abuse; all others chose some issue related to cardiovascular disease. Understanding these differences is a critical first step in aligning priorities for future health promotion activities. As it stands, this information identifies those health centers that share a common health priority with the PHH-PRC and thus make logical partners for initial program planning efforts targeting heart health among these housing development residents.

Readiness scores for the identified health priorities indicate that only 2 of our 15 communities were ready for immediate program implementation (as indicated by a score of 6 or higher). Our assessments indicate that most communities are unprepared to take action despite awareness that there may be a health issue. This suggests that, for future prevention programs to be successful in these communities, there are a number of actions required first to raise awareness and to motivate community members to address the health issue. Information obtained across each community during the readiness assessments provides a rich platform to develop strategies that are specific to each community. For example, key respondents identified specific community members who are considered respected leaders and thus might serve as targets for awareness raising efforts. Immediate next steps in these communities would be meeting with these leaders to get recommendations on the best strategy to raise awareness across the respective community.

The mismatch identified between housing development and health center readiness is not surprising if one considers that health centers exist to address community health. Yet, one might expect that communities that are able to identify their own health priority might be further along the overall continuum of readiness. Given that this is the first application of this model to compare communities, there is no prior literature to draw from in synthesizing our findings. Despite this mismatch between housing development and health center communities, it is interesting that there is consistency across dimensions, even with the observed differences in overall phases of readiness. All communities scored highest in efforts but lowest in knowledge of efforts, suggesting that immediate next steps to program planning need to target education of community leaders and members. Yet the content of this education will differ between housing developments and health centers; according to the stage-specific strategies in the readiness model, targeted messages for housing developments should focus on the importance of heart health, and increasing knowledge of existing programs is needed for staff at health centers.

Analyses of these qualitative data suggest barriers across the housing development communities are largely related to social determinants of health, and confirm the notion that basic life needs preclude action and motivation to address chronic disease despite their high prevalence. These findings are consistent with the vast literature on barriers to accessing needed healthcare in urban, underserved populations.[7,30,31] Our findings also provide objective evidence for community leaders that targeting these barriers when planning prevention programs is essential to effectively impact health outcomes. Thus, our current work in program planning within these communities will address these barriers by adapting a patient navigation model,[32-35] which is designed to target such barriers to accessing health care services.

Our qualitative analyses also indicate that residents of Boston Public Housing Developments have a large variety of clinical care site options, which may contribute to the inertia to engage in care. Because there are so many choices, there is no single clinical setting for this "community" to seek care and knowledge, thus diluting the effect on any programmatic efforts at individual health centers. Although this finding may be unique to the Boston community given the large presence of health care training programs and the network of safety net community health centers, for this partnership it clearly suggests the need for more of a population-level approach for program planning targeting residents of Boston public housing. So, rather than matching developments and health centers on a one-to-one basis, we are planning for housing development navigator training to include comprehensive systems training with all local health centers as well as the common barriers reported by development respondents, and will thus be able to assist residents in accessing care anywhere in Boston.

We found the methods employed in the readiness assessment to be a useful tool for engaging these communities. First, the use of a validated measure to stage level of readiness

is helpful, as community partnership intervention teams often rely on limited impressions or subjective sources. In addition, having multiple respondents and common questions per community allowed for direct comparisons. As a result, the process generated mechanisms for feedback which will help communities make changes.

For example, the study group facilitated discussion of study findings that occurred in multiple community settings. Each community received a personalized report with action-oriented recommendations for their own use. As outlined, these specific findings guided the development of an active community-based patient navigation heart health intervention. In ensuring that all stakeholders were involved from the feasibility phase onward, we were able to avoid potential roadblocks and able to plan a more community appropriate intervention to address heart health within housing developments. In addition, having this platform for engagement served to introduce the research team to future partners and begin the process of building trust and buy-in. In fact, the Boston University Clinical Translational Sciences Institute is planning to use these findings to inform the local research community across both the institution and in the health centers about potential collaborators with common interests for future studies. Although our study has limited generalizability because it was conducted in a single urban setting, our approach utilized common methods and protocols and thus provides a generalizable tool for use in other settings.

CONCLUSIONS

The readiness assessment conducted here provides a foundation for both a dialogue between these two community settings as well as an action-oriented framework to guide collaborations. Eliciting perspectives of key stakeholders (residents of each housing development and staff of each corresponding health center) around a community-identified health priority helped to ensure community buy-in and support for program planning. Future collaborative efforts will be aligned with each community's level of readiness to address their priority health issue to ensure successful implementation. Given the validated methods employed, this approach to program planning may serve as a model for other community collaborations that seek to address a common preventable health issue in their target community.

ACKNOWLEDGMENTS

Our sincere appreciation to both the Boston University Partners in Health in Housing Prevention Research Center (PHH-PRC) and the Boston University Clinical Translational Sciences Institute (BU-BRIDGE) for their commitment to advancing the health of Boston, Massachusetts residents through community-based participatory research. This work would not be possible without the community members who volunteered to discuss the needs of their communities. The authors thank Linda Stanley, from the Tri-Ethnic Center for Prevention Research at Colorado State University, for providing training and guidance on our use of the Community Readiness Model. The findings and conclusions in this article are those of the Partners In Health and Housing Prevention Research Center at Boston University and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

1. Krieger J, Higgins DL. Housing and health: Time again for public health action. Am J Public Health. 2002;92(5):758-68.
2. Victor RG, Leonard D, Hess P, Bhat DG, Jones J, Vaeth PA, et al. Factors associated with hypertension awareness, treatment, and control in Dallas County, Texas. Arch Intern Med. 2008;168(12):1285-93.
3. Smith SC, Jr. Bridging the treatment gap. Am J Cardiol. 2000; 85(12A):3E-7E.
4. Screening for type 2 diabetes mellitus in adults: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2008;148(11):846-54.
5. He J, Muntner P, Chen J, Roccella EJ, Streiffer RH, Whelton PK. Factors associated with hypertension control in the general population of the United States. Arch Intern Med. 2002;162(9):1051-8.

6. Boston Housing Authority [homepage on the Internet; updated 2000; cited 2011 Jul 25]. Available from: http://www.bostonhousing.org/index.html
7. Digenis-Bury EC, Brooks DR, Chen L, Ostrem M, Horsburgh CR. Use of a population-based survey to describe the health of Boston public housing residents. Am J Public Health. 2008; 98(1):85-91.
8. Swider SM. Outcome effectiveness of community health workers: An integrative literature review. Public Health Nurs. 2002; 19(1):11-20.
9. Sikkema KJ, Kelly JA, Winett RA, Solomon LJ, Cargill VA, Roffman RA, et al. Outcomes of a randomized community-level HIV prevention intervention for women living in 18 low-income housing developments. Am J Public Health. 2000;90(1):57-63.
10. Feathers JT, Kieffer EC, Palmisano G, Anderson M, Janz N, Spencer MS, et al. The development, implementation, and process evaluation of the REACH Detroit Partnership's Diabetes Lifestyle Intervention. Diabetes Educ. 2007;33(3):509-20.
11. Vicini F, Nancarrow-Tull J, Shah C, Chmielewski G, Fakhouri M, Sitarek SA, et al. Increasing accrual in cancer clinical trials with a focus on minority enrollment: The William Beaumont Hospital Community Clinical Oncology Program Experience. Cancer. 2011 Mar 31 [Epub ahead of print].
12. Sheikh A, Halani L, Bhopal R, Netuveli G, Partridge MR, Car J, et al. Facilitating the recruitment of minority ethnic people into research: Qualitative case study of South Asians and asthma. PLoS Med. 2009;6(10):e1000148.
13. Baquet CR, Henderson K, Commiskey P, Morrow JN. Clinical trials: The art of enrollment. Semin Oncol Nurs. 2008;24(4):262-9.
14. Partners in Health and Housing Prevention Research Center (PHH-PRC). [homepage on the Internet; updated 2011; cited 2011 Jul 26]. Available from: http://www.bu.edu/phhprc/contact-us/
15. Rorie JA, Smith A, Evans T, Horsburgh CR, Jr., Brooks DR, Goodman R, et al. Using resident health advocates to improve public health screening and follow-up among public housing residents, Boston, 2007-2008. Prev Chronic Dis. 2011; 8(1):A15.
16. Plested BA, Edwards RW, Jumper-Thurman P. Community readiness: A handbook for successful change. Fort Collins (CO): Tri-Ethnic Center for Prevention Research; 2011.
17. Edwards RW, Jumper-Thurman P, Plested BA, Oetting ER, Littlethunder L. Community readiness: Research to practice. J Community Psychol. 2000;28(3):291-307.
18. Oetting ER, Donnermeyer JF, Plested BA, Edwards RW, Kelly K, Beauvais F. Assessing community readiness for prevention. Int J Addict. 1995;30(6):659-83.
19. Oetting ER, Jumper-Thurman P, Plested B, Edwards RW. Community readiness and health services. Subst Use Misuse. 2001;36(6-7):825-43.
20. Peercy M, Gray J, Thurman PJ, Plested B. Community readiness: An effective model for tribal engagement in prevention of cardiovascular disease. Fam Community Health. 2010;33(3):238-47.
21. Findholt N. Application of the Community Readiness Model for childhood obesity prevention. Public Health Nurs. 2007; 24(6):565-70.
22. Sliwa S, Goldberg JP, Clark V, Collins J, Edwards R, Hyatt RR, et al. Using the Community Readiness Model to select communities for a community-wide obesity prevention intervention. Prev Chronic Dis. 2011;8(6):A150.
23. Vernon IS, Jumper-Thurman P. Prevention of HIV/AIDS in Native American communities: Promising interventions. Public Health Rep 2002;117(Suppl 1):S96-103.
24. Lawsin CR, Borrayo EA, Edwards R, Belloso C. Community readiness to promote Latinas' participation in breast cancer prevention clinical trials. Health Soc Care Community. 2007; 15(4):369-78.
25. Israel BA, Schulz AJ, Parker EA, Becker AB. Community-based participatory research: Policy recommendations for promoting a partnership approach in health research. Educ Health (Abingdon). 2001;14(2):182-97.
26. Boston Public Housing Authority. April 2011 Demographic Report. Boston: Author; 2011. (Internal Report).
27. Boston University Clinical and Translational Science Institute. Boston HealthNet Community Health Center Data [updated 2008]. Available from: http://www.ctsi.bu.edu/index.php/community/community-health-centers/.com
28. Boyatzis R. Transforming qualitative information: Thematic analysis and code development. Thousand Oaks (CA): Sage; 1998.
29. Schroepfer TA, Waltz A, Noh H, Matloub J, Kue V. Seeking to bridge two cultures: The Wisconsin Hmong cancer experience. J Cancer Educ. 25(4):609-16.
30. Bisgaier J, Rhodes KV. Cumulative adverse financial circumstances: Associations with patient health status and behaviors. Health Soc Work. 2011;36(2):129-37.
31. Black HL, Priolo C, Akinyemi D, Gonzalez R, Jackson DS, Garcia L, et al. Clearing clinical barriers: Enhancing social support using a patient navigator for asthma care. J Asthma. 2010;47(8):913-9.
32. Wells KJ, Battaglia TA, Dudley DJ, Garcia R, Greene A, Calhoun E, et al. Patient navigation: State of the art or is it science? Cancer. 2008;113(8):1999-2010.
33. Paskett ED, Harrop JP, Wells KJ. Patient navigation: An update on the state of the science. CA Cancer J Clin. 2011;61(4):237-49.
34. Battaglia TA, Roloff K, Posner MA, Freund KM. Improving follow-up to abnormal breast cancer screening in an urban population. A patient navigation intervention. Cancer. 2007; 109(2 Suppl):359-67.
35. Freeman HP, Rodriguez RL. History and principles of patient navigation. Cancer. 2011;117(15 Suppl):3539-42.

DOI:10.1158/1055-9965.EPI-12-0532

Boston Patient Navigation Research Program: The Impact of Navigation on Time to Diagnostic Resolution after Abnormal Cancer Screening

Tracy A. Battaglia[1], Sharon M. Bak[1], Timothy Heeren[2], Clara A. Chen[3], Richard Kalish[4], Stephen Tringale[5], James O. Taylor[6], Barbara Lottero[7], A. Patrick Egan[8], Nisha Thakrar[9], and Karen M. Freund[1,10]

Abstract

Background: There is a need for controlled studies to assess the impact of patient navigation in vulnerable cancer populations.

Methods: Boston Patient Navigation Research Program conducted a quasi-experimental patient navigation intervention across six federally qualified inner-city community health centers, three assigned to a breast cancer navigation intervention and three assigned to a cervical cancer navigation intervention; each group then served as the control for the other. Eligible women had an abnormal breast or cervical cancer screening test conducted at one of the participating health centers during a baseline (2004–2005) or intervention period (2007–2008). Kaplan–Meier survival curves and proportional hazards regression examined the effect of patient navigation on time to definitive diagnosis, adjusting for covariates, clustering by clinic and differences between the baseline and intervention period.

Results: We enrolled 997 subjects in the baseline period and 3,041 subjects during the intervention period, of whom 1,497 were in the navigated arm, and 1,544 in the control arm. There was a significant decrease in time to diagnosis for subjects in the navigated group compared with controls among those with a cervical screening abnormality [aHR 1.46; 95% confidence interval (CI), 1.1–1.9]; and among those with a breast cancer screening abnormality that resolved after 60 days (aHR 1.40; 95% CI, 1.1–1.9), with no differences before 60 days.

Conclusions: This study documents a benefit of patient navigation on time to diagnosis among a racially/ethnically diverse inner city population.

Impact: Patient navigation may address cancer health disparities by reducing time to diagnosis following an abnormal cancer-screening event. *Cancer Epidemiol Biomarkers Prev; 21(10); 1645–54. ©2012 AACR.*

Introduction

There is widespread recognition that increasing complexity of cancer care contributes to outcome disparities (1, 2) as evidenced by the impact that lack of access to timely, quality cancer services has on outcomes among vulnerable populations defined by low income, inadequate insurance coverage, and minority race/ethnicity (3–7). To address this failure to equally deliver the spectrum of cancer services to all Americans, there is growing emphasis on patient navigation programs (8–10).

Patient navigation was designed to identify and eliminate barriers to accessing cancer services to achieve timely completion of care (11, 12). The goal of patient navigation is to facilitate timely care by meeting cultural needs and addressing barriers to accessing care at the individual level. Examples of navigation services include identifying patients at risk for delays, facilitating appointment scheduling, coordinating care among providers, organizing interpreter services, ensuring access to prior medical records, and providing logistic support for transportation, childcare, or linkage to community resources. Navigators are trained to help patients advocate for themselves in the health care system and provide emotional support during this stressful period.

Several demonstration projects and small clinical trials report the potential benefit of patient navigation at certain points in the cancer care continuum and in specific populations (10, 13–15) resulting in several organizations recommending patient navigation as standard care (16, 17). However, large-scale clinical trials across diverse

Authors' Affiliations: [1]Women's Health Unit, Section of General Internal Medicine, Evans Department of Medicine, Boston Medical Center and Women's Health Interdisciplinary Research Center, Boston University School of Medicine; [2]Department of Biostatistics; and [3]Data Coordinating Center, Boston University School of Public Health; [4]BostonHealthNet; [5]Codman Square Community Health Center; [6]East Boston Neighborhood Health Center; [7]Greater Roslindale Medical and Dental Center; [8]Dorchester House Community Health Center; [9]South Boston Community Health Center; and [10]Institute for Clinical Research and Health Policy Studies, Tufts Medical Center and Tuft University School of Medicine, Boston, Massachusetts

Corresponding Author: Tracy A. Battaglia, Boston University School of Medicine, 801 Massachusetts Avenue, Crosstown 1st Floor, Boston, MA 02118. Phone: 617-638-8036; Fax: 617-638-8096; E-mail: Tracy.Battaglia@bmc.org

doi: 10.1158/1055-9965.EPI-12-0532

DOI:10.1158/1055-9965.EPI-12-0532

populations are still lacking. Implementation research is needed to assess the ability of navigation to be adopted in usual clinical practice, especially in safety net institutions that care for low-income patients (18). To provide controlled clinical trial data on the efficacy and effectiveness of patient navigation for follow-up of screening abnormalities, the Center to Reduce Cancer Health Disparities (CRCHD) and the American Cancer Society funded 10 intervention sites through 9 grants (11) to conduct a national Patient Navigation Research Program (PNRP). The Boston PNRP, one of the funded programs, was designed as an effectiveness study to assess the impact of patient navigation under usual clinical practice rather than the ideal circumstances of a randomized clinical trial.

Materials and Methods

Study design

We conducted a quasi-experimental intervention across 6 federally qualified inner-city community health centers targeting women with breast and cervical cancer screening abnormalities. The Boston program was designed as a clinical effectiveness study of patient navigation as a new standard of care, where all patients with screening abnormalities were included into the trial and thus the Boston University Institutional Review Board approved the study with a waiver of written informed consent. Using the community based participatory research process (19, 20) guided by a Community Advisory Panel, each health center partner agreed to collaborate in developing and evaluating the patient navigation program at their site. Each health center was assigned a navigation site for either abnormal breast or cervical cancer screening and a control for the other condition. To adjust for temporal trends and observe for contamination between the intervention and control conditions, we collected baseline data from the health centers on a random sample of all subjects with abnormal screening in 2004 to 2005, and collected data on all subjects with abnormal screening during the intervention period from January 2007 to December 2008. We followed all subjects for 365 days, censoring those who had not achieved diagnostic resolution by the end of the study.

Study sites

Because there were only 6 sites, each serving different racial/ethnic communities, rather than randomly assigning each to the navigation condition, the study team developed a strategic allocation protocol to reduce imbalance across the study by race/ethnicity.

Inclusion and exclusion criteria

We used electronic medical records to identify eligible women with an abnormal breast or cervical cancer screening test conducted at one of the participating health centers during the study periods. Breast abnormalities include abnormal mammograms, ultrasound or magnetic resonance imaging (MRI) of the breasts, coded as BI-RADS 0, 3, 4, or 5 using the Breast Imaging Reporting and Data System (21, 22), or abnormal clinical breast examinations with a mass or other lesion suspicious for cancer. Cervical cancer screening abnormalities included Pap test findings of low- and high-grade squamous intraepithelial lesions (LGSIL and HGSIL), atypical cells of unknown significance (ASCUS), when reflex testing for high risk human papilloma virus (HPV) was positive, or atypical glandular cells of unknown significance (AGUS). Women were excluded if they were pregnant at the time of the screening abnormality, as pregnancy may alter the evaluation protocol and its time course, or had a cognitive impairment that prevented their participation with a patient navigator. During the baseline period, we sampled all high-grade (BI-RADS 4, 5, and HGSIL) lesions and a random sample of all other mammographic and cytologic lesions with up to 100 subjects per health center (23). We included all women with abnormalities during the intervention period. When guidelines for management of ASCUS HPV–positive and LGSIL changed for women ages 18 to 21 years in 2007 (24), recommending annual Pap tests and not diagnostic colposcopy, we excluded, thereafter, those subjects where the clinician choose repeat Pap tests as the management, retaining any subjects if the clinician referred the patient for colposcopy.

Recruitment, training, and supervision of patient navigators

Patient navigators were hired as employees of the health centers, not part of the research staff. With personnel turnover, there were a total of 23 female navigators. They all had a high school education and some health care experience, such as medical assistant training. None had advanced training in social work or nursing. Each health center attempted to hire navigators who could address the language needs of their communities; five navigators spoke another language, including Spanish, Vietnamese, Albanian, and Portuguese Creole; all other navigators had access to robust interpreter services. They were integrated into the clinical team at the health center, receiving direct supervision from health center staff. All received standard training to ensure minimal competency. The research team conducted bi-monthly local training for all navigators and their supervisors throughout the study period. All navigators participated in annual face-to-face national PNRP trainings and several webinars to standardize the training across the national sites (25). A protocol was developed at each health center for identifying women with abnormal cancer screening. Navigators contacted patients by telephone to initiate navigation after a clinician had informed the patient of the test results. Navigators used the care management model (26) to identify barriers to recommended care, and developed strategies to address these barriers, with the focus on timely completion of the diagnostic evaluation. Follow-up occurred by telephone, mail, and in face-to-face meetings, usually at the health center. Navigators documented their activities in a structured template in the electronic medical

DOI:10.1158/1055-9965.EPI-12-0532

record, which allowed providers to view the activities in real time and allowed electronic capture of the navigator activities.

The research team conducted annual core competency assessments on all navigators, and provided feedback to the local supervisors. The research team also provided each site with a monthly spreadsheet of patients eligible for navigation at the health center and their progress in reaching diagnostic resolution.

Data collection

All data were captured through the electronic medical records at each of the health centers regardless of control or navigation status. Demographic information was captured from registration and billing systems. Electronic charts recorded all screening studies and their results, so we were able to capture abnormalities electronically. All subsequent information on the clinical diagnostic evaluation, including tests ordered, completed, and their results, were collected through manual abstraction by trained abstractors, with 10% of charts reviewed by a second abstractor for accuracy of abstraction with 3% errors identified.

Race/ethnicity was collected from registration data, as our prior work indicated a high correlation of this administrative data with patient self-report (27). We imputed the $n = 229$ (8%) missing race/ethnicity data from the following sources, in order of: providers' notes, country of origin (28), last name (29), or based on the majority race/ethnicity group of their community health center site. Missing language ($n = 174$, 6%) was imputed as other than English if there was documentation in the medical record of the primary language, of need for interpreter, or correspondence with the patient in another language.

Outcomes

The primary outcomes of interest in our study were whether and when diagnostic resolution of the screening abnormality was achieved, as defined by the Design and Analysis Committee for the National PNRP (11). Time to diagnostic resolution was defined as the time from the date of the initial screening abnormality to the date when the final definitive diagnostic test or evaluation was completed, a timepoint determined by the clinicians caring for the patient. For BI-RADS 4 and 5 lesions, it was usually biopsy. For BI-RADS 3 imaging, where serial 6 month imaging for up to 2 years is recommended, we observed only to confirm when the next imaging was completed after 6 months. We subtracted 180 days from the diagnostic evaluation time, which allowed us to compare all abnormalities in a similar timeframe. If the BI-RADS 3 was addressed in less than 180 days, we used 0.5 days for the time to diagnostic resolution. For BI-RADS 0 lesions, resolution was defined as either follow-up imaging designated BI-RADS 1 or 2, or completion of whatever diagnostic testing through biopsy had been recommended. For cervical abnormalities, diagnostic resolution was usually colposcopy.

Analysis

We followed our *a priori* analysis plan to conduct a difference in differences survival analysis, comparing the differences across the baseline and intervention time periods between the navigated and control study sites. We calculated separate Kaplan–Meier survival curves and proportional hazards regression models for breast and cervical subjects to examine the effect of patient navigation on time to definitive diagnosis for each clinical abnormality. Proportional hazards regression models examined the effect of patient navigation, controlling for differences between the baseline and intervention periods, grade of abnormality, and demographic covariates. We included as potential confounders the covariates of age (breast categories: 18–40, 41–64, >65 years and cervical categories: 18–20, 21–30, >30 years), high (BIRADS 4 and 5, HGSIL) versus lower (BIRADS 0 and 3, LGSIL) grade screening abnormality, race/ethnicity, primary language and insurance (private, public or no insurance), and percentage of zip code under federal poverty level as proxy for socioeconomic status. These models accounted for clustering by health center. We included interactions terms with the time to diagnostic resolution variable where proportional hazards assumptions were violated.

All analyses were conducted in SAS, and the proportional hazards regression procedure in SAS used a robust sandwich estimator for the covariance matrix in Wald tests for the global null hypothesis and individual parameters to account for clustering (30). The effect of patient navigation met assumptions of proportionality for the cervical cancer screening arm of the trial, but did not for the breast cancer screening arm. We noted that the survival curves for the baseline and intervention period for the breast navigated sites crossed at approximately 60 days, and therefore included an interaction term between intervention and time allowing differential effects of patient navigation over days 0 to 59 and days 60 to 365. All observations were censored at 365 days. A limitation of the sandwich estimator to account for clustering is that it is biased downward when there are few clusters (31). Therefore we conducted a secondary analysis using a shared gamma frailty model with a proportional hazards model fit through a penalized likelihood approach (using the frailtypack procedure in R; 32) as a second approach to account for clustering. This method does not allow one to account for the nonproportional hazards for navigation in the breast cancer screening arm in a single model. Therefore, we ran separate proportional hazards for diagnostic resolution within 59 days and for diagnostic resolution between 60 and 365 days in this secondary analysis.

Because we were specifically interested in knowing whether navigation had a greater impact on vulnerable groups of women, described here by race/ethnicity, primary language, and insurance coverage, we developed models including interaction terms for these variables with the navigation variable. Our primary analysis was conducted as an intention to treat analysis, including all women whether or not they received patient navigation.

DOI:10.1158/1055-9965.EPI-12-0532

We conducted a secondary analysis, excluding those navigated subjects where the navigator did not initiate patient navigation within 30 days for breast screening abnormalities and 60 days for cervical screening abnormalities.

Results

There were a total of 997 subjects in the baseline period and 3,041 subjects during the intervention period, of which 1,497 received navigation and 1,544 were controls. Tables 1 and 2 display demographics separately for breast and cervical subjects, looking at the overall group divided by both study period (baseline versus intervention) and by study site (navigation versus control sites). The majority of subjects were from racial or ethnic minority groups, including 30% from African American communities and 29% from Hispanic populations. The subjects spoke 18 primary languages, with Spanish (19%), Vietnamese (3.5%), Albanian (2%), and Portuguese (1%) being the most common. The differences in the proportion of women from each racial and ethnic group in the navigated versus control sites reflects the different communities served by each health center, and different sizes of the health centers. There are also differences in race/ethnicity over time within the health center populations. The proportion uninsured dropped from 40% to 29% overall between the baseline and intervention time periods, reflecting the impact of Massachusetts Health Insurance Reform, initiated in 2006 (33, 34).

Figure 1 represents the unadjusted proportion of subjects who ever reach a diagnostic resolution by 365 days, comparing differences across the 2 study periods (baseline and intervention) within navigation and control sites separated for breast and cervical cancer subjects. For the breast cancer screening arm, at baseline, the control sites had similar rates of diagnostic resolution compared with the navigated sites (89.6% vs. 89.7%). The control sites showed no improvement in the proportion of women reaching diagnostic resolution from the baseline to the intervention period (89.6% vs. 90.2%, $P = 0.7809$), whereas the percentage of women at navigated sites who resolved increased from 89.7% to 92.6% from the baseline to

Table 1. Demographic characteristics of breast cancer screening subjects in the Boston Patient Navigation Research Program by Study Site Group (navigation vs. control sites) across the 2 study periods (baseline vs. intervention; $N = 2,275$)

	Baseline period 2004–2005			Intervention period 2007–2008			
	Navigation sites, *N* (%)	Control sites, *N* (%)	*P*[a]	Navigation sites, *N* (%)	Control sites, *N* (%)	*P*[a]	All sites, *N* (%)
Breast screening abnormalities	312	211		772	980		2,275
Age							
18–40	29 (9)	12 (6)	0.29	154 (20)	88 (9)	<0.0001	283 (12)
41–64	239 (77)	171 (81)		540 (70)	755 (77)		1,705 (75)
65+	44 (14)	28 (13)		78 (10)	137 (14)		287 (13)
Insurance							
Uninsured	112 (36)	87 (41)	0.09	137 (18)	346 (35)	<0.0001	682 (30)
Public	108 (35)	54 (26)		375 (48)	357 (36)		894 (39)
Private	92 (29)	70 (33)		260 (34)	277 (29)		699 (31)
Race/ethnicity							
African American	74 (24)	105 (50)	<0.0001	124 (16)	349 (36)	<0.0001	652 (29)
Hispanic	66 (21)	36 (17)		160 (21)	327 (33)		589 (26)
White	145 (46)	58 (27)		346 (45)	290 (30)		839 (37)
Other[b]	27 (9)	12 (6)		142 (18)	14 (1)		195 (8)
Language							
English	205 (67)	145 (70)	0.05	490 (63)	545 (56)	<0.0001	1,385 (61)
Spanish	57 (19)	24 (12)		104 (13)	258 (26)		443 (20)
Other[c]	42 (14)	39 (18)		178 (23)	177 (18)		436 (19)
BI-RADS category							
BIRAD 0	235 (75)	117 (55)	<0.0001	494 (64)	801 (82)	<0.0001	1,647 (72)
BIRAD 3	59 (19)	71 (34)		82 (11)	144 (15)		356 (16)
BIRAD 4,5	18 (6)	23 (11)		10 (1)	15 (1)		66 (3)
Clinical breast exam	0	0		186 (24)	20 (2)		206 (9)

[a]*P* value for test for χ^2 test.

[b]Most common response was Asian (of which 71% were Vietnamese).

[c]Most common other languages included Vietnamese (25%), Albanian (14%), Portuguese, and Portuguese Creole (5%).

DOI:10.1158/1055-9965.EPI-12-0532

Table 2. Demographic characteristics of cervical cancer screening subjects in the Boston Patient Navigation Research Program by Study Site Group (navigation vs. control sites) across the 2 study periods (baseline vs. intervention; $N = 1{,}763$)

	Baseline period 2004–2005			Intervention period 2007–2008			
	Navigation sites, *N* (%)	Control sites, *N* (%)	*P*[a]	Navigation sites, *N* (%)	Control sites, *N* (%)	*P*[a]	All sites, *N* (%)
Cervical screening abnormalities	234	240		725	564		1,763
Age, y							
18–20	31 (13)	17 (7)	0.001	82 (11)	67 (12)	0.04	197 (11)
21–30	114 (49)	158 (66)		389 (54)	337 (60)		998 (57)
30+	89 (38)	65 (27)		254 (35)	160 (28)		568 (32)
Insurance							
None	109 (46)	87 (37)	0.09	287 (40)	107 (19)	<0.0001	590 (34)
Public	72 (31)	77 (33)		281 (39)	221 (39)		651 (37)
Private	53 (23)	69 (30)		157 (21)	236 (42)		515 (29)
Race/ethnicity							
African American	105 (45)	55 (23)	<0.0001	285 (39)	123 (22)	<0.0001	568 (32)
Hispanic	52 (22)	77 (32)		304 (42)	117 (21)		550 (31)
White	33 (14)	88 (37)		129 (18)	277 (49)		527 (30)
Other[b]	44 (19)	20 (8)		7 (1)	47 (8)		118 (7)
Language							
English	163 (71)	158 (78)	0.09	447 (62)	478 (85)	<0.0001	1,246 (72)
Spanish	38 (17)	33 (16)		215 (30)	39 (7)		325 (19)
Other[c]	27 (12)	12 (6)		63 (8)	47 (8)		149 (9)
Cervical abnormality							
Low grade[d]	205 (88)	209 (87)	0.86	685 (96)	526 (95)	0.27	1,625 (93)
High grade[e]	29 (12)	31 (13)		27 (4)	28 (5)		115 (7)

[a]*P* value for test of homogeneity of proportions.
[b]Most common response was Asian (of which 53% were Vietnamese).
[c]Most common other languages included Vietnamese (20%), Albanian (8%), Portuguese, and Portuguese Creole (19%).
[d]Includes the following Pap test results: ASCUS/HPV+ (Atypical squamous cells of undetermined significance/positive for human papillomavirus) and LGSIL (low-grade squamous intraepithelial lesion).
[e]Includes the following Pap test results: AGUS, HGSIL, and carcinoma.

intervention period, although this difference did not reach statistical significance ($P = 0.12$). For the cervical cancer screening arm, the control sites had slightly higher rates of diagnostic resolution during the baseline period compared with the navigated sites (87.9% vs. 79.1%). The control sites again showed no change temporally at 365 days after abnormal screening (80.0% versus 78.6%, $P = 0.64$). However, the navigated sites showed significant improvement in diagnostic resolution comparing the baseline with the intervention period (79.1% vs. 87.9%; $P = 0.0008$).

Figures 2 and 3 display the Kaplan–Meier curves for time to diagnostic resolution (where the *y*-axis represents the proportion achieving diagnostic resolution) comparing the baseline and intervention periods across both the navigation and control sites. For breast cancer screening subjects in the navigation sites, there is a superimposition of the baseline and intervention period curves through about 60 days, when 73% of subjects in the baseline period have achieved resolution compared with 71% of subjects in the intervention period. Median days to resolution are also similar in the 2 periods, with median days to resolution of 29 days in the baseline period and 32 days in the intervention period. After 60 days, there is a consistently higher rate of resolution among subjects during the intervention than the baseline period. For patients without resolution by 60 days, median days to resolution decreased from 157 days to 70 days from the baseline to intervention period. The control sites had no change between the baseline and intervention time periods, also reflected at 60 days when 73% of subjects in the baseline period achieved resolution compared with 72% of subjects in the intervention period, and with median days to resolution of 27 days in the baseline period and 34 days in the intervention period. For cervical cancer screening subjects in the navigation sites, the intervention period showed more subjects achieved timely resolution almost immediately. For example, at 60 days, only 27% of subjects in the baseline period achieved diagnostic resolution compared with 39% of subjects in the intervention period,

DOI:10.1158/1055-9965.EPI-12-0532

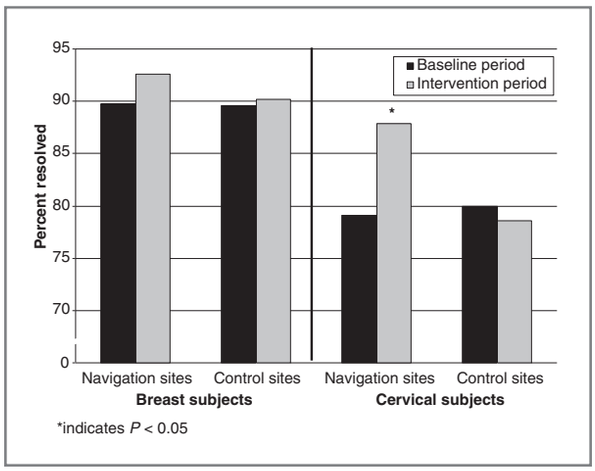


Figure 1. Percent of subjects in the Boston Patient Navigation Research Program who complete diagnostic resolution by 365 days by Study Period (baseline vs. intervention) within each Study Site Group (navigation vs. control sites), for both breast and cervical subjects.

and median days to resolution decreased from 110 in the baseline period to 76 in the intervention period. These improvements among intervention subjects in the cervical navigation sites persisted throughout the 365 day follow-up, compared with the baseline period, and compared with the control sites in either time period where the proportion of subjects achieving diagnostic resolution at 60 days was similar for baseline and intervention subjects (37% vs. 31%, respectively), and median time to resolution was also similar (84 days vs. 90 days).

Table 3 presents the adjusted HRs for the proportional hazards models, for breast and cervical cancer subjects, respectively. We found a significant benefit of patient navigation in reaching diagnostic resolution after adjusting for differences in differences between the navigation and control sites of the study and the intervention and baseline time periods. The model also adjusted for covariates and for clustering within each of the health centers. In the breast model, the adjusted HR for completing diagnostic evaluation in the first 59 days is 1.04 (95% CI, 0.83–1.30), indicating no difference between the navigated and control sites after adjustment. From days 60 through 365, there was a significant impact of patient navigation (aHR 1.40; 95% CI, 1.06–1.86), where HR more than 1.0 indicates that patient navigation is associated with more timely diagnostic resolution. For the cervical cancer screening subjects of the study, the adjusted HR was 1.46 (95% CI, 1.14–1.88), indicating a benefit of navigation over the control sites in reaching diagnostic resolution throughout the course of the study. Secondary analyses using a shared gamma frailty model to account for clustering gave similar effect size for breast cancer screening subjects, but were not significant in separate models for the first 59 days (aHR 1.03; 95% CI, 0.8–1.3) or for days 60–365 (aHR 1.5; 95% CI, 0.9–2.4), whereas significance remained among the cervical cancer screening subjects (aHR 1.4; 95% CI, 1.1–1.8).

In the breast model, several covariates were associated with delays in achieving diagnostic resolution. Compared with patients who had private insurance, those with public insurance (aHR 0.8; 95% CI, 0.8–0.9), but not uninsured (aHR 0.9; 95% CI, 0.8–1.0; $P = 0.07$), were less likely to have timely resolution. Patients with a primary language other than English or Spanish were also less likely to have timely resolution (aHR 0.8; 95% CI, 0.7–0.9). Index screening abnormalities of BI-RADS 3 (aHR 12.7; 95% CI, 8.7–18.6) or BI-RADS 4 or 5 (aHR 4.5; 95% CI, 2.0–10.0) were initially associated with more timely resolution, which decayed over time, compared with those with a

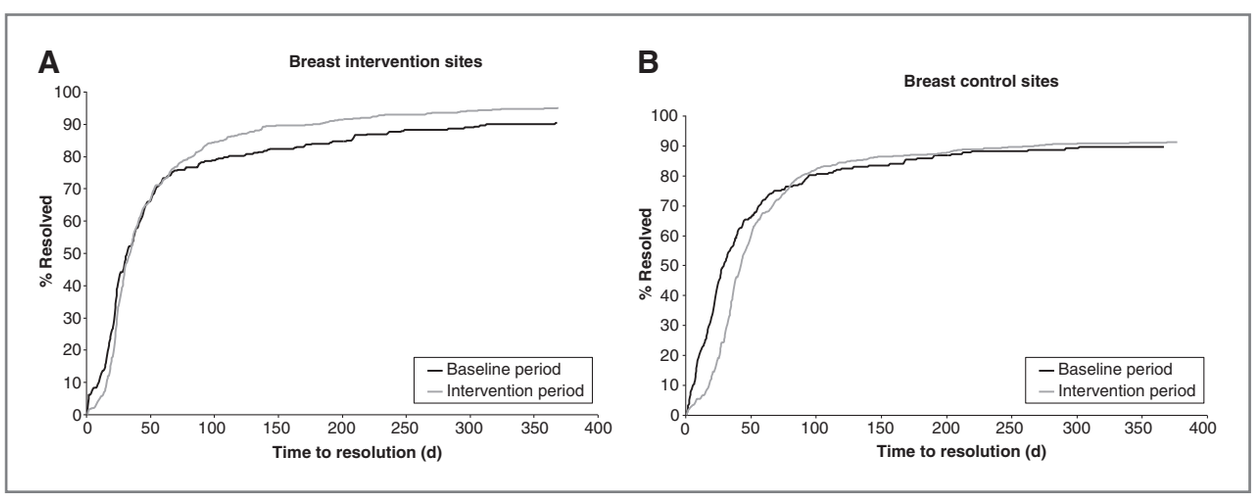


Figure 2. Kaplan–Meier survival curves of time to diagnostic resolution of breast cancer screening subjects in the Boston Patient Navigation Research Program comparing baseline and intervention periods across both navigation and control study sites. A, survival curve of time to diagnosis at breast cancer screening navigation sites, before (baseline period, 2004–5) and during (intervention period 2007–8) implementation of the navigation intervention. B, survival curve of time to diagnosis at Breast cancer screening control sites, before (baseline period 2004–5) and during (intervention period 2007–8) implementation of the navigation intervention.

DOI:10.1158/1055-9965.EPI-12-0532

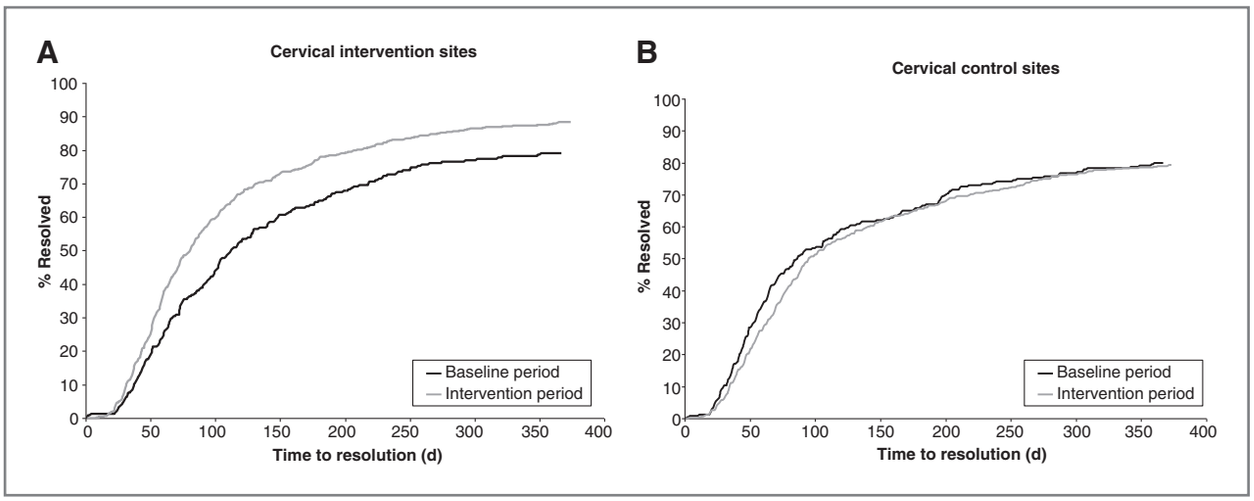


Figure 3. Kaplan–Meier survival curves of time to diagnostic resolution of cervical cancer screening subjects in the Boston Patient Navigation Research Program comparing baseline and intervention periods across both navigation and control study sites. A, survival curve of time to diagnosis at cervical cancer screening navigation sites, before (baseline period, 2004–5) and during (intervention period, 2007–8) implementation of the navigation intervention. B, survival curve of time to diagnosis in cervical cancer screening control sites, before (baseline period, 2004–5) and during (intervention period, 2007–8) implementation of the navigation intervention.

BI-RADS 0 screening abnormality. Age, race/ethnicity, and percentage of zip code below federal poverty level were not predictors of timely resolution in the model.

In the cervical model, public insurance status and younger age were associated with delays in achieving diagnostic resolution. Compared with private insurance, publically insured (aHR 0.8; 95% CI, 0.7–0.9) but not uninsured women (aHR 1.0; 95% CI, 0.9–1.2) were less likely to have timely resolution. Women ages 18 to 40 (aHR 0.7; 95% CI, 0.6–0.9) were less likely to have timely resolution than older women. Race/ethnicity, language, screening abnormality, and socioeconomic status were not predictors of the outcome.

We were specifically interested in examining whether vulnerable populations defined by minority race/ethnicity, language other than English, or public or no insurance had a benefit from the intervention. None of the interaction terms of these variables by the navigation variable were significant in either the breast or cervical model, suggesting that there was no differential benefit of navigation for any specific group of women defined by these demographic variables.

Our data indicated that 124 (16.1%) women in the navigated arm did not receive navigation services within 30 days of the breast screening abnormality, and 103 (14.2%) of women did not receive navigation within 60 days of the cervical cancer screening abnormality. These findings varied by health center, and ranged from 4.0% to 29.6%. Removing women enrolled later into navigation did not change our findings (breast navigation effect $\geq$ 60 days: aHR, 1.4; 95% CI, 1.0–1.8; cervical navigation effect: aHR, 1.6; 95% CI, 1.3–2.1). Median time that a navigator spent per case varied by disease status; breast navigators spent a median of 60 minutes per subject with navigation activities [75% interquartile range (IQR), 30–127.5], with a median of 1 encounter per subject (range 0–15); whereas cervical navigators spent a median of 75 minutes (75% IQR, 30–120) in navigation activities, with a median of 2 encounters per subject (range 0–18).

Discussion

The Boston Patient Navigation Research Program showed a benefit of patient navigation for vulnerable communities of women with both breast and cervical cancer screening abnormalities with an adjusted HR between 1.4 and 1.5. Observing the navigated arm over 365 days, this translated into an additional 9% of patients with cervical abnormalities and 3% of women with breast screening abnormalities completing diagnostic resolution to determine whether or not they have a premalignant or invasive cancer. By including all individuals with a screening abnormality, this study represents an assessment of the effectiveness of the patient navigation intervention for an entire population cared for in a safety net system of care, leading to greater external validity. Individual randomized clinical trial methodologies are often unable to randomize the most difficult to reach patients who may be in greatest need of navigation (14), and thus limits their external validity. Our effectiveness design includes this difficult to reach group of women while ensuring identical methods in the navigated and control arms of the study, and thus provides an estimate of the effect of navigation when applied to an entire practice population. Our data provide scientifically rigorous evidence of how navigation can be implemented.

The generalizability of our findings extends to a diverse group of women from multiple communities and cultures: more than 55% of our sample were from minority communities, and more than 60% were either publicly insured or uninsured. Each community health center that

DOI:10.1158/1055-9965.EPI-12-0532

Table 3. Cox proportional adjusted HRs (aHR) for time to resolution, according to type of cancer screening in subjects with an abnormal breast ($N = 2,275$) or cervical ($N = 1,763$) cancer screening exam during intervention time period compared with baseline time period (Boston Patient Navigation Research Program)

Screening cancer	Adjusted HR (95% CI)[a]	*P*
Breast		
Control sites	1.0 (reference)	
Navigation sites		
Resolution before 60 days	1.04 (0.83–1.30)	0.746
Resolution after 60 days	1.40 (1.06–1.86)	0.019
Cervical		
Control sites	1.0 (reference)	
Navigation sites	1.46 (1.14–1.88)	0.003

[a]Cox proportional analysis adjusted for age, race/ethnicity, language, insurance, index abnormality, and socioeconomic status, and for clustering with community health center.

participated in the trial provided care to unique immigrant and minority communities in the greater Boston area. Our work shows that navigation is broadly beneficial to all groups. We did not find that our navigation intervention preferentially supported any group of subjects, defined by race/ethnicity, language, or insurance status. The increased rates of insurance seen in our intervention period of 2007 to 08 compared with the baseline 2004 to 05 periods reflect the impact of Massachusetts Health Insurance reform. The persistent uninsured rates of 8% to 47% by health center, all higher than the overall 5% uninsured in the state in 2008 (35), reflect that women who are uninsured despite insurance reform disproportionately seek their care at community health centers. Our finding that publically insured subjects had less timely care compared with uninsured subjects is consistent with other studies and suggests that insurance status may be a marker for other social determinants of health (36, 37).

The benefit of patient navigation does not seem to be only from shifting the diagnostic completion curve to an earlier time for all, but also in increasing the proportion completing diagnostic resolution. Our data in the breast cancer screening abnormalities suggests that navigation does not show an immediate benefit, but may be of greatest benefit for those with initial delays. This may be because many patients who complete in a timely manner will do so without the help of a navigator. This has implications for institutions with limited resources that wish to develop navigation programs with the most effective impact. Several navigation programs addressing annual mammography screening, for example, began at 18 months, providing patients a 6-month window to complete screening on their own before activating navigation (38).

The main limitation of our trial is the lack of randomization of subjects into the intervention and control groups. This could have resulted in potential imbalance of both measured and unmeasured differences between the groups. We attempted to address this design limitation by collecting baseline data from all sites. This allowed us to adjust for secular trends and baseline differences in the health centers. The lack of any differences over time in the control groups suggest that there was no secular trend to account for our findings. It also suggests that there was little contamination of the navigation intervention into the control arms. A related limitation is that our study was conducted at 6 urban, East Coast community health centers. While our analyses do account for clustering of patients within health centers, available methods of accounting for clustering in the context of proportional hazards regression for time-to-event data do not perform well with a small numbers of clusters. Our results may not generalize to less urban settings, or other regions of the country, and we acknowledge that our data collection methods did not account for care received outside of the affiliated health care system, though our previous work has shown that the majority of this population have documented longevity within this safety-net health system (39, 40).

To successfully disseminate patient navigation programs, there are several components critical to ensure success. Safety-net institutions, including community health centers and hospital ambulatory centers, will require an initial and ongoing training program to ensure the competence of navigators. Care sites will require the ability to track patients with screening abnormalities, and link this tracking to the work of the navigators. The methods developed for the health centers in this project fulfill these components while also meeting many of the principles of the Patient Centered Medical Home, including criteria for care coordination, registry/tracking, and meaningful use of electronic medical records (41). Because this study was not powered to provide clarity on all aspects of navigation, including its impact among specific sub-populations or on patient-reported outcomes or cost-effectiveness, it lays the groundwork for the future studies necessary to inform best practices.

This study highlights the benefits of a community based participatory research process (20) where the design of the research protocol as a quality improvement effectiveness research project was developed with leadership of the health centers and under the guidance of an active Community Advisory Panel. The findings provide controlled clinical trial data on the effectiveness of patient navigation to promote equal access to timely diagnostic cancer care for a racial/ethnically diverse low-income population. Perhaps the strongest indicator of the success of this intervention is the fact that the health centers continued to implement the intervention with internal funds following completion of the trial and saw this as a key

DOI:10.1158/1055-9965.EPI-12-0532

component of their Patient Centered Medical Home implementation.

Disclosure of Potential Conflicts of Interest

K.M. Freund has a commercial research grant from Komen for the Cure and Avon Foundation and is a consultant/advisory board member of Komen for the Cure and Avon Foundation. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: T.A. Battaglia, R. Kalish, K.M. Freund
Development of methodology: T.A. Battaglia, S.M. Bak, S. Tringale, K.M. Freund
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): T.A. Battaglia, S.M. Bak, S. Tringale, J.O. Taylor, A.P. Egan, K.M. Freund
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): T.A. Battaglia, S.M. Bak, T.C. Heeren, C.A. Chen, R. Kalish, K.M. Freund
Writing, review, and/or revision of the manuscript: T.A. Battaglia, S.M. Bak, T.C. Heeren, C.A. Chen, R. Kalish, B. Lottero, A.P. Egan, K.M. Freund
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): T.A. Battaglia, S.M. Bak, K.M. Freund
Study supervision: T.A. Battaglia, S.M. Bak, K.M. Freund
Other (study and staff supervision at community health center): N. Thakrar

Grant Support

This work was funded by the National Cancer Institute and the Center to Reduce Cancer Health Disparities, National Cancer Institute (U01 CA116892-01).

Received May 1, 2012; revised August 20, 2012; accepted August 21, 2012; published online October 8, 2012.

References

1. American Cancer Society. Cancer facts & figures 2011. Atlanta, GA: American Cancer Society; 2011.
2. Smedley BD, Stith AY, Nelson AR. Unequal treatment: confronting racial and ethnic disparities. Washington, DC: The National Academies Press; 2003.
3. Kohler BA, Ward E, McCarthy BJ, Schymura MJ, Ries LA, Eheman C, et al. Annual report to the nation on the status of cancer, 1975–2007, featuring tumors of the brain and other nervous system. J Natl Cancer Inst 2011;103:714–36.
4. Eheman C, Henley SJ, Ballard-Barbash R, Jacobs EJ, Schymura MJ, Noone AM, et al. Annual Report to the Nation on the status of cancer, 1975–2008, featuring cancers associated with excess weight and lack of sufficient physical activity. Cancer 2012;118:2338–66.
5. Howe HL, Wu X, Ries LA, Cokkinides V, Ahmed F, Jemal A, et al. Annual report to the nation on the status of cancer, 1975–2003, featuring cancer among U.S. Hispanic/Latino populations. Cancer 2006;107:1711–42.
6. Weir HK, Thun MJ, Hankey BF, Ries LA, Howe HL, Wingo PA, et al. Annual report to the nation on the status of cancer, 1975–2000, featuring the uses of surveillance data for cancer prevention and control. J Natl Cancer Inst 2003;95:1276–99.
7. Shavers VL, Fagan P, McDonald P. Health disparities across the cancer continuum. J Health Care Poor Underserved 2007;18:1–5.
8. Battaglia TA, Burhansstipanov L, Murrell SS, Dwyer AJ, Caron SE. Assessing the impact of patient navigation: prevention and early detection metrics. Cancer 2011;117:3553–64.
9. Dohan D, Schrag D. Using navigators to improve care of underserved patients: current practices and approaches. Cancer 2005;104:848–55.
10. Wells KJ, Battaglia TA, Dudley DJ, Garcia R, Greene A, Calhoun E, et al. Patient navigation: state of the art or is it science? Cancer 2008; 113:1999–2010.
11. Freund KM, Battaglia TA, Calhoun E, Dudley DJ, Fiscella K, Paskett E, et al. National Cancer Institute Patient Navigation Research Program: methods, protocol, and measures. Cancer 2008;113:3391–9.
12. Freeman HP, Rodriguez RL. History and principles of patient navigation. Cancer 2011;117:3539–42.
13. Battaglia TA, Roloff K, Posner MA, Freund KM. Improving follow-up to abnormal breast cancer screening in an urban population. A patient navigation intervention. Cancer 2007;109:359–67.
14. Ell K, Vourlekis B, Lee PJ, Xie B. Patient navigation and case management following an abnormal mammogram: a randomized clinical trial. Prev Med 2007;44:26–33.
15. Paskett ED, Harrop JP, Wells KJ. Patient navigation: an update on the state of the science. CA Cancer J Clin 2011;61:237–49.
16. American College of Surgeons. Commission on Cancer. Cancer Program Standards 2012: Ensuring patient centered care. Available from: https://www.sgmc.org/sites/www/Uploads/COC%20Program%20Stds%202012.pdf
17. American College of Surgeons. National accreditation program for breast centers; 2011 [cited 2011 Mar 28]. Available from: http://accreditedbreastcenters.org/accreditation/application.html.
18. Freund KM. Patient navigation: the promise to reduce health disparities. J Gen Intern Med 2010;26:110–2.
19. Ramsden VR, McKay S, Crowe J. The pursuit of excellence: engaging the community in participatory health research. Glob Health Promot 2011;17:32–42.
20. Israel BA, Schulz AJ, Parker EA, Becker AB. Community-based participatory research: policy recommendations for promoting a partnership approach in health research. Educ Health (Abingdon) 2001;14: 182–97.
21. American College of Radiology (ACR). Illustrated breast imaging reporting and data System (BI-RADS). 4th ed. Reston VA: American College of Radiology; 2003.
22. Obenauer S, Hermann KP, Grabbe E. Applications and literature review of the BI-RADS classification. Eur Radiol 2005;15:1027–36.
23. Battaglia TA, Santana MC, Bak S, Gokhale M, Lash TL, Ash AS, et al. Predictors of timely follow-up after abnormal cancer screening among women seeking care at urban community health centers. Cancer 2010;116:913–21.
24. Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. Am J Obstet Gynecol 2007;197:346–55.
25. Calhoun EA, Whitley EM, Esparza A, Ness E, Greene A, Garcia R, et al. A national patient navigator training program. Health Promot Pract 2009;11:205–15.
26. Longest BB, Young GJ. Coordination and communication. In: Shortell SK, Kaluzny AD, editors. Health care management: organizational design and behavior. 4th ed. Albany, NY: Delmar Publishers; 2000. p. 237–75.
27. Laiteerapong N, Sherman BJ, Kim K, Freund KM. Validation of racial categorization in a hospital administrative database. J Gen Intern Med 2005;20:130.
28. Schenker N, Parker JD. From single-race reporting to multiple-race reporting: using imputation methods to bridge the transition. Stat Med 2003;22:1571–87.
29. Word DL, Perkins RC. Building a Spanish Surname List for the 1990's - a new approach to an old problem. Technical Working Paper No. 13 March 1996. Population Division, U.S. Bureau of the Census: Washington DC, 1996.
30. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. J Am Stat Assoc 1989;4:1074–8.
31. Murray DM, Varnell SP, Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. Am J Public Health 2004;94:423–32.

DOI:10.1158/1055-9965.EPI-12-0532

32. Rondeau V, Gonzalez JR. frailtypack: a computer program for the analysis of correlated failure time data using penalized likelihood estimation. Comput Methods Programs Biomed 2005;80:154–64.
33. McDonough JE, Rosman B, Phelps F, Shannon M. The third wave of Massachusetts health care access reform. Health Aff (Millwood) 2006;25:w420–31.
34. Pande AH, Ross-Degnan D, Zaslavsky AM, Salomon JA. Effects of healthcare reforms on coverage, access, and disparities: quasi-experimental analysis of evidence from Massachusetts. Am J Prev Med 2011;41:1–8.
35. Long SK, Stockley K. Sustaining health reform in a recession: an update on Massachusetts as of fall 2009. Health Aff (Millwood) 2010;29:1234–41.
36. Ayanian JZ, Kohler BA, Abe T, Epstein AM. The relation between health insurance coverage and clinical outcomes among women with breast cancer. N Engl J Med 1993;329:326–31.
37. Roetzheim RG, Pal N, Tennant C, Voti L, Ayanian JZ, Schwabe A, et al. Effects of health insurance and race on early detection of cancer. J Natl Cancer Inst 1999;91:1409–15.
38. Phillips CE, Rothstein JD, Beaver K, Sherman BJ, Freund KM, Battaglia TA. Patient navigation to increase mammography screening among inner city women. J Gen Intern Med 2010;26:123–9.
39. Freund K, Hanchate A, Isabelle A, Kalish R, Kapoor A, Bak S, et al. The impact of insurance instability on care after abnormal cancer screening. J Gen Int Med 2012;27(S1):S333.
40. Peterson NB, Han J, Freund KM. Inadequate follow-up for abnormal Pap smears in an urban population. J Natl Med Assoc 2003;95:825–32.
41. Center for Medicare and Medicaid Services (CMS). Electronic health records and meaningful use. The Office of the National Coordinator for Health Information Technology; 2011 [cited 2012 Apr 9]. Available from: http://healthit.hhs.gov/portal/server.pt?open=512&objID=2996&mode=2.

Prevention of Fatal Opioid Overdose

Leo Beletsky, JD, MPH
Josiah D. Rich, MD, MPH
Alexander Y. Walley, MD, MSc

OPIOID OVERDOSE IS A BURGEONING PUBLIC HEALTH crisis, accounting for at least 16 000 deaths annually in the United States.[1] Opioid overdose occurs across sex, ethnic, age, and geographic strata and involves both medical and nonmedical opioid use. To date, federal government response has focused primarily on monitoring and securing the drug supply.[1] This Viewpoint suggests various steps necessary to support a more comprehensive approach.

During the time it typically takes some overdoses to turn fatal, it is possible to reverse the respiratory depression and other effects of opioids with the antagonist naloxone. Community-based organizations, health care institutions, and local and state agencies have begun to train and equip potential nonmedical bystanders to recognize and reverse overdose events using first aid techniques and emergency supplies of naloxone.[2,3] As the number of such initiatives has increased, the 53 000 program trainees have tracked more than 10 000 reports of overdose rescues in the United States.[2] These efforts have targeted drug users (syringe access programs, drug treatment centers, correctional facilities),[1] physicians (to "coprescribe" naloxone along with opioids),[4] and first responders (ie, fire and police).[3] The concept has also gained traction among policy makers, including the Office of the National Drug Control Policy and professional organizations.[5]

Despite the mounting supportive evidence, the number of these programs remains limited in many communities with elevated rates of fatal overdose.[2] Multiple barriers limit the diffusion of this innovation: the price of naloxone has skyrocketed in the context of a severe shortage; few prescribers are aware of and are willing to facilitate overdose prevention education and naloxone access; funding for program activities and evaluation research remains sparse; and the Food and Drug Administration (FDA)–approved formulation of naloxone is suboptimal for out-of-hospital use.

In April 2012, an interagency hearing on naloxone access was convened by the FDA and brought together practitioners, regulators, researchers, and people personally affected by overdose. Although the meeting underscored numerous regulatory hurdles, decisive action is necessary to advance overdose prevention programs beyond the proof-of-concept phase (TABLE).

The FDA must ensure that adequate supplies of naloxone are available to meet the increasing demand. Like many sterile injectable products, naloxone is in chronic shortage. Most naloxone programs experience challenges in obtaining naloxone because of cost increases or suppliers' inability to fill orders.[1] As existing programs scale up and more jurisdictions adopt these measures, federal action can help expand naloxone supply, such as by fast-tracking importation licenses.

Health care practitioners are optimally positioned to facilitate opioid overdose prevention. Equipping clinicians and providing incentive for them to screen patients for overdose risk and to educate patients, their families, and caregivers about recognizing and responding to overdoses is an important step, alongside community-based education about opioid overdose and naloxone distribution. Explicitly integrating overdose prevention, including naloxone coprescription, into the FDA-industry cooperative strategy for evaluation and mitigation of opioid risk also could be helpful.[4]

Clinicians may be unclear about legal risks associated with prescribing naloxone[6] and may be concerned about the possibility of facilitating risky drug use[6]; yet there is no evidence of such disinhibition.[7] Prescribing naloxone to manage opioid overdose is consistent with its FDA-approved indication, precipitating no increased liability as long as prescribers adhere to general rules of professional conduct.[7] Some states have passed laws indemnifying clinicians from risk of malpractice lawsuits perceived to arise from prescription of naloxone. Others have introduced Good Samaritan laws shielding lay bystanders and persons experiencing overdose from possible civil liability (flowing from providing first aid) and criminal drug charges when 911 is called.[7] Using evidence-based model legislation, federal coordination can help disseminate these legal protections to encourage clinician engagement, lay responder rescue, and help-seeking.

In out-of-hospital settings, the administration of injectable drugs carries the risk of needle-stick injury and presents logistical barriers, such as the absence of a sterile syringe and delay in preparation.[3] FDA action is needed to fast-track approval of naloxone delivery systems that are safe and user-friendly for nonmedical responders. Administration of intranasal naloxone via aftermarket nasal atomizers is an "off-label" system increasingly used by out-of-hospital emergency medical personnel and by community-based programs.[3] The lack of FDA approval limits the implementation of intranasal formulations and devices: nasal atomizers are difficult to stock and seldom covered by insurance. FDA approval of intranasal naloxone is predicated on research demonstrating such a formulation to be "substantially

Author Affiliations: Northeastern University School of Law and Bouvé College of Health Sciences, Boston, Massachusetts (Mr Beletsky); University of California, San Diego, School of Medicine, La Jolla (Mr Beletsky); Brown Medical School and The Miriam Hospital, Providence, Rhode Island (Dr Rich); and Boston University School of Medicine and Boston Medical Center, Boston (Dr Walley).
Corresponding Author: Leo Beletsky, JD, MPH, Northeastern University School of Law and Bouvé College of Health Sciences, Boston, Massachusetts, 400 Huntington Ave, Boston, MA 02115 (l.beletsky@neu.edu).

Table. Potential Federal Action to Improve Prevention of Fatal Overdose

Barrier/Limiting Factor	Remedial Federal Action	Agency in Charge
Naloxone shortage and high cost	Allow temporary importation from qualifying foreign manufacturers	FDA
	Monitor supply and demand	FDA
	Provide funding for dissemination of overdose prevention and management education, including naloxone distribution to potential bystanders through community-based and addiction treatment programs	SAMHSA, CDC
Lack of health care training to screen for and address problematic opioid use, opioid overdose risk, and underutilization of out-of-hospital naloxone	Require Opioid REMS programs to cover overdose prevention and the role of naloxone	FDA
	Create and disseminate existing practitioner training and toolkits (such as the materials available at http://prescribetoprevent.org/) to facilitate opioid overdose prevention Materials should be tailored to priority settings and populations, including chronic pain therapy, drug treatment facilities, and emergency departments	SAMHSA, AHRQ, national professional associations
Lack of prescriber incentives to deliver overdose prevention and management education	Reimburse overdose prevention education interventions and cover naloxone prescription and delivery devices	CMS
Prescriber concern about legal risks to self and patients from naloxone prescription; bystander reluctance to call 911	Formulate and disseminate model legislation providing legal immunity to naloxone prescribers and lay responders to change police behaviors and encourage witnesses to call 911	DOJ, Bureau of Justice Assistance
Lack of public awareness about risk factors, signs and symptoms, and appropriate response to overdose	Conduct public awareness campaigns about risk factors, signs and symptoms, and response to overdose	FDA, CDC, ONDCP
Research	Evaluate community-based naloxone access initiatives, Good Samaritan laws, and interventions targeting drug users and possible bystanders on overdose risk behaviors	NIH, CDC, DOJ
	Evaluate the effects of education, brief interventions, laws and regulations, and other interventions on clinician practice and patient overdose morbidity and mortality outcomes	NIH, CDC, DOJ
	Provide funding for program evaluation and clinical research on intranasal and autoinjector devices	NIH, CDC, SAMHS
Explore alternative models of naloxone access	Such models may include a new "over-the-counter-plus" class of drugs and pharmacist naloxone prescription	FDA, national professional associations

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; CDC, Centers for Disease Control and Prevention; CMS, Centers for Medicare & Medicaid Services; DOJ, Department of Justice; FDA, Food and Drug Administration; NIH, National Institutes of Health; ONDCP, Office of the National Drug Control Policy; SAMHSA, Substance Abuse and Mental Health Services Administration.

equivalent" to the current formulation. Federal funding is needed to clear this hurdle because naloxone is an off-patent, generic medication not widely considered to be a promising investment by major pharmaceutical companies. Use of an autoinjector—a safe and effective method for delivery of epinephrine—represents another alternative that would require FDA approval.

Federal agencies are uniquely situated to address national public health crises through increasing awareness, funding, and coordinated action. Community-based organizations, state and local health departments, and professional societies have taken the lead in developing, implementing, and publicizing overdose education and naloxone distribution as a component in a comprehensive response to this veritable epidemic.[5] The federal government should actively support research and programmatic action on overdose education and naloxone access.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: Mr Beletsky is supported by award R01DA01982905 from the National Institute on Drug Abuse (S. Strathdee, principal investigator [PI]). Dr Rich is supported by awards 5K24DA022112 and 1R21DA029201 from the National Institute on Drug Abuse. Dr Walley is the Medical Director of the Massachusetts Department of Public Health Opioid Overdose Prevention Pilot Program and is supported by awards R25-DA13582 and R01DA032082 (J. Samet, PI) from the National Institute on Drug Abuse and award U01AA11003 (R. Saitz, PI) from the National Institute on Alcohol Abuse and Alcoholism.

Role of the Sponsors: The sponsors had no role in design, preparation, review, approval, or decision to publish this article.

Disclaimer: The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Drug Abuse, the National Institute on Alcohol Abuse and Alcoholism, the National Institutes of Health, the Massachusetts Department of Public Health, or the Substance Abuse and Mental Health Services Administration.

REFERENCES

1. Centers for Disease Control and Prevention (CDC). CDC grand rounds: prescription drug overdoses—a U.S. epidemic. *MMWR Morb Mortal Wkly Rep*. 2012; 61(1):10-13.
2. Centers for Disease Control and Prevention (CDC). Community-based opioid overdose prevention programs providing naloxone—United States, 2010. *MMWR Morb Mortal Wkly Rep*. 2012;61(6):101-105.
3. Doe-Simkins M, Walley AY, Epstein A, Moyer P. Saved by the nose: bystander-administered intranasal naloxone hydrochloride for opioid overdose. *Am J Public Health*. 2009;99(5):788-791.
4. Albert S, Brason FW II, Sanford CK, et al. Project Lazarus: community-based overdose prevention in rural North Carolina. *Pain Med*. 2011;12(s2)(suppl 2):S77-S85.
5. AMA Adopts New Policies at Annual Meeting. American Medical Association website. http://www.ama-assn.org/ama/pub/news/news/2012-06-19-ama-adopts-new-policies.page. June 19, 2012. Accessed June 20, 2012.
6. Beletsky L, Ruthazer R, Macalino GE, et al. Physicians' knowledge of and willingness to prescribe naloxone to reverse accidental opiate overdose. *J Urban Health*. 2007;84(1):126-136.
7. Burris S, Beletsky L, Castagna C, et al. Stopping an invisible epidemic: legal issues in the provision of naloxone to prevent opioid overdose. *Drexel Law Rev*. 2009; 1(2):273-340.

J Nurs Care Qual
Vol. 27, No. 3, pp. 258–265

Team Improvement and Patient Safety Conferences

Culture Change and Slowing the Revolving Door Between Skilled Nursing Facility and the Hospital

Randi E. Berkowitz, MD; Robert Schreiber, MD; Michael K. Paasche-Orlow, MD, MA, MPH

To improve the safety culture of a skilled nursing facility, we conducted multidisciplinary "Team Improvement for Patient and Safety" (TIPS) case conferences biweekly to identify causes of transfers to acute care hospitals and improvement opportunities. Staff perceptions of organizational patient safety culture were assessed with the Nursing Home Survey on Patient Safety Culture. Over the course of the year, we held 22 TIPS conferences. Mean item scores increased during the study, indicating improved staff perceptions of patient safety culture ($P < .005$). **Key words:** *multidisciplinary conferences*, *nursing home*, *patient safety*, *quality improvement*, *safety culture*, *team conferences*

NURSING HOME PATIENTS are frequently admitted to hospitals.[1-3] This is costly and contributes to poor patient outcomes. In addition, a proportion of these events are avoidable, as has been demonstrated by several recentlypublished interventions.[1,4] Of 1.8 million skilled nursing facility (SNF) admissions in the United States during 2006, 23.5% resulted in readmission to an acute care hospital within 30 days at a cost of $4.3 billion.[3] Unplanned transfers from SNFs for acute patient problems are often associated with adverse events or medication errors and are burdensome to the patient and family.[1] Known causes of acute care hospital transfers from the SNF include lack of on-site medical staff, unavailability of rapid laboratory testing services or immediate intravenous fluids and medications, inadequate assessment to identify clinical deterioration, communication gaps, and lack of clarity regarding advance directives.[4,5]

Several interventions have been shown to successfully decrease 30-day rehospitalization rates in the acute care setting. Some projects have involved various approaches for activating patients, families, and providers to improve patient safety during care transitions. For example, Coleman et al[6] showed that care transition coaches decrease rehospitalizations even up to 180 days after hospital discharge. Similarly, Naylor et al[7] showed that advanced practice nurses can lower hospital

Author Affiliations: *Hebrew SeniorLife (Drs Berkowitz and Schreiber); Division of Gerontology, Beth Israel Deaconess Medical Center, Harvard Medical School (Drs Berkowitz and Schreiber); and Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine (Dr Paasche-Orlow), Boston, Massachusetts.*

The authors thank Eric Coleman, Kyle Allen, Richard DellaPenna, Alice Bonner, and Rich Jones for their help on this project. This study was funded by Practice Change Fellowship, Denver, Colorado.

The authors declare no conflict of interest.

Correspondence: *Randi Berkowitz, MD, 1200 Centre St, Roslindale, MA 02459 (rberkowitz@brca.harvard.edu).*

Accepted for publication: January 28, 2012

Published online before print: February 22, 2012.

DOI: 10.1097/NCQ.0b013e31824623a4

readmissions in patients with congestive heart failure. Jack et al[8] used inpatient nurses with a checklist and personalized patient education materials along with a postdischarge pharmacist telephone call, which resulted in a one-third decrease in the rate patients returned to the hospital within 30 days.

Few interventions, however, have focused on transitions from the SNF to the hospital.[9,10] Ouslander et al[1] using a comprehensive rehospitalization reduction program called INTERACT and the Evercare demonstration project[11] have decreased hospital transfers from nursing facilities. More recently, Berkowitz et al[12] described a successful intervention to improve transitions of care in an SNF. This 3-pronged intervention included (1) an admission checklist for physicians to encourage discussion of repeat hospitalizations before worsening of symptoms, (2) automatic palliative care consultations for patients with more than 3 rehospitalizations, and (3) Team Improvement for Patient and Safety (TIPS) conferences, recurring multidisciplinary team meetings to analyze root causes of failures in care.

The TIPS conferences were designed to increase the organizational commitment to patient safety and transitions of care among frontline staff. This was motivated, in part, by studies that have shown that a higher degree of organizational commitment to patient safety, as reflected by the Nursing Home Survey on Patient Safety Culture, is linked to improved patient outcomes in SNFs.[13,14] In this article, we describe in detail the design and implementation of a successful TIPS program that was part of the larger intervention resulting in a 20% reduction in transfer rates on our skilled nursing unit.[12] Staff attitudes on patient safety culture before and during the intervention are discussed, based on biannual responses to the benchmarked Nursing Home Survey on Patient Safety Culture designed specifically for nursing home providers and staff; the survey evaluates opinions regarding the culture of patient safety in a nursing home.[15]

METHODS

Nursing facility

The project was implemented on a 50-bed subacute rehabilitation unit admitting approximately 1000 patients per year. The medical staff members were all salaried employees consisting of 1.5 full-time equivalent of a nurse practitioner and 1.5 full-time equivalent of physician time. The unit is housed in a 600-unit long-term care, religious-affiliated, not-for-profit organization in Boston, Massachusetts. In addition to standard rehabilitation care, on-site services include radiology, electrocardiograph (ECG) technician, mobile laboratory, blood transfusion, psychiatry, geriatric psychiatry, and a palliative care team composed of a chaplain, a nurse, a physician, and a social worker.

Planning process

The project was approved by the Hebrew SeniorLife institutional review board as a quality improvement project. A multidisciplinary team was convened to understand the scope of the problem, determine objectives for the project, develop concrete steps to address the problem of avoidable acute care hospital transfers, and reassess outcomes. The team was composed of the medical director, the director of nursing, the administrator, staff nurses and aides, therapists, a rabbi from the palliative care team, a daughter whose mother had experienced multiple rehospitalizations, a statistician, and representatives from a local acute care hospital and a home health agency. This group met twice before the project began to determine the format of TIPS conferences, approach to root-cause analyses of care transitions, outcome measures, and how to involve the palliative care team.

Measures

To measure fidelity to the project and give credit to those who participated, attendance of core staff was recorded at TIPS meetings. Core staff included nurses, physicians, nurse practitioners, administrators, nursing administrators, a member of the quality

improvement committee, therapists, social workers, and palliative care team members.

Staff attitudes on the organizational patient safety culture were monitored at baseline and biannually with the Nursing Home Survey on Patient Safety Culture.[15] The surveys were distributed to the core target audience for the TIPS meeting. Surveys were sent by e-mail, and a paper version also was available. The Nursing Home Survey on Patient Safety Culture includes 42 survey items measuring 12 dimensions and an additional overall resident safety question. The dimensions are as follows: teamwork (4 items), staffing (4 items), compliance with procedures (3 items), training and skills (3 items), nonpunitive response to mistakes (4 items), handoffs (4 items), feedback and communication about incidents (4 items), communication openness (3 items), supervisor expectations and actions promoting resident safety (3 items), overall perceptions of resident safety (3 items), management support for resident safety (3 items), and organizational learning (4 items.) Answers to all Likert scale questions were coded 1 to 5, ranging from "strongly disagree" to "strongly agree." In addition, an overall rating of patient safety in the facility was included as the 43rd item.

Mean composite dimensions scores for the perception of organizational patient safety culture were calculated by summing the scores for individual questions (with appropriate inversion of reverse-coded items) and dividing by the number of questions answered. Respondents who answered "does not apply/don't know" were treated in the same way as those who did not answer the item (missing). In addition, following Agency for Healthcare Research and Quality methods, the mean percentage of positive responses was derived by scoring each question answered "strongly agree" or "agree" as a positive response. The mean percentage of positive responses overall was calculated by adding the number of respondents who reported "agree" or "strongly agree" divided by the number of respondents for that item. The percentage of positive responses was calculated both for the entire survey and for 7 specific dimensions (teamwork, nonpunitive response to mistakes, feedback and communication about incidents, supervisor expectation and actions promoting resident safety, organizational learning, and communication openness) of the survey identified by the project team to be most relevant to the TIPS program.

Statistical analysis

Single-factor analyses of variance were calculated to compare the (1) mean overall composite scores, (2) mean scores for the 7 specified dimensions thought to be most relevant to the intervention, (3) mean percentage of positive responses, and (4) mean percentage of positive responses for composite dimensions—across the 3 time points of survey administration (ie, preinitiation of TIPS conferences, 6 months, and 12 months). Data were analyzed with Microsoft Excel 2003.

Intervention

TIPS conferences were conducted every 2 weeks during the 1-year study period. All direct care staff members were invited and were told that attendance would be taken. Other staff as well as patients and their families were invited (eg, ECG technicians, maintenance, security, information technology, home care agency staff, an outpatient pharmacist, family members, and patients) on a case-by-case basis. Cases of potentially avoidable acute care hospital transfers or adverse events that might have ended in an acute care hospital transfer were identified by the medical director, hospital nursing liaisons, the nursing quality improvement specialist, members of the palliative care team, patients, families, and staff, as well as providers at acute care hospitals and home care agencies, They were all asked to report such information to the medical director (Table). Potential cases were preliminarily reviewed by the medical director and scheduled for a TIPS conference session within 1 to 2 weeks of the event at a time key participants could attend. The TIPS conference sessions were designed to last 30 minutes. In

Table. Examples of New Procedures to Reduce Rehospitalization and Improve Patient Care Developed From TIPS Conferences

1. Standardized discharge checklist (includes patient and family education and sending discharge summary to outpatient pharmacy, home care agency, and primary care provider)
2. Protocols to reduce risk of medication errors
   a. Anticoagulation protocol checklist with pharmacy, nursing, and provider daily review
   b. Sign on medication carts, "Medication pass in progress," to minimize interruptions/distractions of nurses during medication administration
   c. Medication list at patient bedside to involve patient and family in detecting potential errors
   d. Medication reconciliation process with patient, family, pharmacy, nursing, and providers
3. Procedures for working with disruptive families
   a. Behavioral plans initiated within 1 d of admission and set expectations for respect shown to staff
   b. Early plan of care for daily patient needs
   c. Call to hospital and outside providers to ensure message similar to families among all providers
4. Procedures for telephone communication with families
   a. Single number at facility for family to call for issues
   b. Care manager staffing this call-in line responsible to correct problem or find appropriate team members to answer concerns
   c. Call-in line for patients who are in the facility or after discharge
5. Procedures related to customer service
   a. Development of plan for patients at appointments to get their next meal
   b. Call light response rate program initiated with hourly rounds
6. Universal transfer form development with our state Department of Public Health

Abbreviation: TIPS, Team Improvement and Patient Safety.

the first 5 minutes, the case was presented. After this, the medical director facilitated a 15-minute group discussion, in which an explicit expectation was set for all attendees to participate. The final 10 minutes of the session was dedicated to the development of action steps. If a patient or family member presented the case, the presenter was excused after the presentation to allow for open staff discussion. The medical director led the majority of the conferences.

Mechanics of the conferences

To encourage development of actionable steps, the leader used a white board to record problems and help the group identify concrete solutions. The leader also made an effort to engage all the staff members in the discussion. Staff members were called on if they were not speaking during the conferences.

To ensure participation by nursing aides and direct care staff, meetings were conducted on the unit to minimize the time staff members were unable to answer call lights and dispense medications. A rotating nurse and aide covered call lights during each conference. In addition, meeting times were varied to allow night and evening shift staff members to attend.

To ensure that the lessons learned from the conferences were broadcast widely to staff both within and outside our organization, an e-mail listserve was created and short "tips from TIPS" were e-mailed on the same day of the conference. This enabled staff members from all shifts who might not have attended a particular conference to learn from the cases as well. The evening supervisor used these e-mails to review the cases with staff on weekends and nights. We invited multiple patients and family members to come back and give

their perspective on outcomes that had not been ideal.

RESULTS

Over the course of the year, we held 22 TIPS conferences. The cases discussed in the TIPS conferences were all held within 2 weeks of the sentinel events. TIPS conference attendance by core unit staff was 81%. TIPS conference discussions led directly to multiple procedural changes to improve patient safety and decrease readmission risk (Table). Baseline, 6-month, and 1-year Nursing Home Survey on Patient Safety Culture response rates were 12.5%, 56.9%, and 49.4%, respectively. Mean scores increased during the study, indicating improved patient safety culture (pre-TIPS, mean [M] = 3.3, n = 10; 6-month, M = 3.5, n = 41; 1-year, M = 3.9, n = 40; analysis of variance, $F_{2,88} = 5.66$, $P < .005$). Percentage of positive response rates for the overall survey and the 7 selected dimensions showed increasing trends during the course of the study (Figure).

DISCUSSION

The TIPS conference initiative improved the staff perception of the patient safety culture of the unit. This likely contributed to the observed 20% decrease in the acute care hospital transfer rates that we have previously reported.[12] Several key aspects that were important to the success of this program were institutional support and program leadership. We had intended to hold TIPS conferences every 2 weeks throughout the course of the year. While we did not meet that goal, holding 22 of 26 (85%) conferences with high attendance represents a high degree of fidelity to the program. The support of the institution's senior management was critical to help build a culture of accountability that was not focused on blame. It also was critical to have a committed program leader to choose appropriate cases, encourage attendance, and ensure implementation of the solutions identified. For example, the SNF had a medication error in which a patient did not receive some doses of medicine. There was an immediate apology to the patient that was publicized broadly within our institution, and the patient presented her experience after this event at a TIPS conference. Pharmacy, nursing, and medical procedures were changed to review the list of all orders for medications that were expiring that day. The chief executive officer of our organization publicly congratulated the team and its efforts. We had several conferences after this in which nurses who themselves had been involved in medication errors requested TIPS conferences to address these events.

Culture change

When the conferences first began, staff nurses did not want to attend the conferences, as they feared that they alone would be blamed. In addition, there was fear that if more acutely ill patients were kept on the unit instead of being sent to the hospital, then the nursing workload would increase, which would adversely impact the care being given to other patients.

To encourage open conversation, the first cases chosen were ones in which the medical director bore the major responsibility for the outcome. Open conversations by the leader about ways in which she might have prevented a transfer helped to foster a nonblaming atmosphere. Thought leaders among direct nursing staff were identified. Once they engaged productively with the TIPS program, they encouraged participation by their peers. If the conversation did shift to blaming, the discussion was redirected by the leader. Aides in particular were reticent to speak in the first conferences, but over time, they began to feel more comfortable voicing concerns.

We anticipated that keeping potentially more acutely ill patients would present challenges for direct care staff. For example, we reviewed the case of a dying woman whose family had decided against a transfer to the hospital and the patient died on our unit on comfort measures. However, the woman required a significant amount of both nursing and aide time to ensure her comfort. In

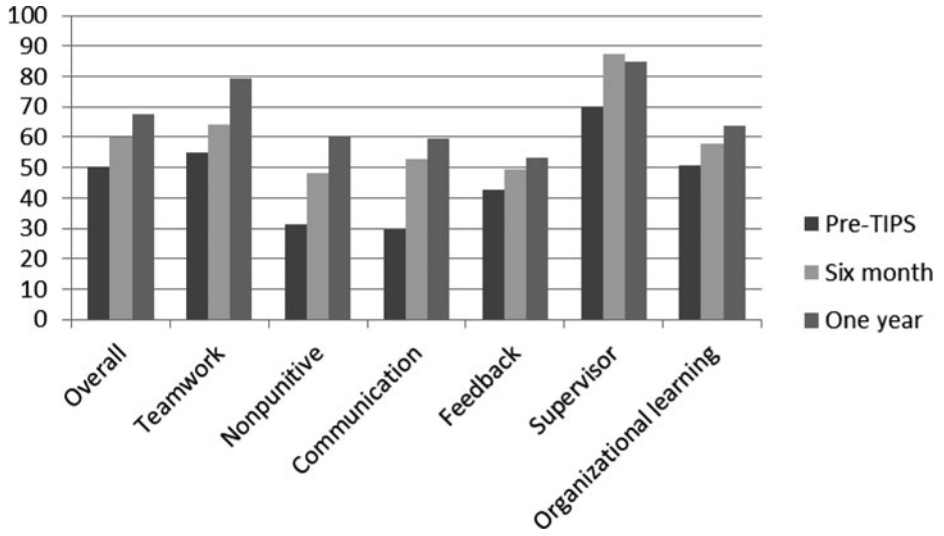


Figure. Percentage of staff reporting "agree" or "strongly agree" from the Nursing Home Survey on Patient Safety Culture to questions on dimensions of patient safety. TIPS indicates Team Improvement and Patient Safety.

response to this case, we brought senior nursing leadership into the TIPS meeting to discuss staffing concerns associated with very ill patients and developed a plan for future cases to use other staff resources in the building such as a nursing supervisor whom staff had been hesitant to call.

Involvement of the entire team

Early on we recognized that many of the issues involved team members both within and outside our core team. In one case, the husband of a dying patient would walk up and down the hallway in the evenings intimidating direct care staff and threatening to sue them. This issue was identified, and through TIPS, a behavioral management plan was implemented. In another case, a woman went out to the hospital with atrial fibrillation because the ECG printout was lost. The ECG technician explained that the machine stored information and could easily have reprinted a hardcopy. She and the secretarial staff developed a process to ensure that the ECG would reliably reach the clinicians. Another TIPS case involved a patient who had gained 20 lb on the unit without this being recognized by the staff. Information technology staff members were brought in to demonstrate how biometric data such as weight could be recorded on our electronic medical record to more effectively track weight changes.

We also involved participants from outside our institution. We reviewed the case of a male patient discharged home from the unit who soon ended up back in the hospital because of a lack of ostomy bags. We had nurses from a home care agency present the case and the downstream effects of not sending bags home with the patient. Outpatient pharmacists were brought in to TIPS conferences as we identified issues of patients who had incorrect medications from outdated medication lists delivered to their homes.

Another benefit of implementing TIPS was the rapid review and discussion of the sentinel event, generally within a week. This increased staff ability to remember details of the events and quickly give a sense of closure around difficult situations. The conferences helped staff to see problems looming early on in a patient's stay. Staff members were frequently heard saying, "We'd better address this before it becomes a TIPS case."

The staff survey results were reviewed, and a subcommittee composed of frontline aides, nurses, therapists, our administrator, and the director of nursing was formed to tackle issues noted in the survey. Some of the presentations at the conferences by patients and families were very moving and helped focus staff members on the patient perspective. Lessons learned were in multiple areas of care delivery: time-out for high-risk discharges, standardized handoff to next team, shift-to-shift communication, easy access for patients and families after discharge, patient and family understanding of illness and prognosis, optimizing delivery of bad news, addressing disruptive families within a day of admission, communication with outside providers for ensuring unified messages to families and review of care, and identifying community resources. Care processes on the SNF unit that were addressed included infection control, medication errors, and customer service.

The project described in this article was 1 element of a 3-pronged approach to improve discharge that also included standardized physician admission assessments and automatic palliative care consultations for high-risk patients. It is not possible to determine which changes in our care processes produced the reduction in hospital admissions or improvement in our patient safety culture. The study was limited by small sample size. Because of this, we were not able to stratify our results by provider class. In addition, the baseline survey included a smaller proportion of participants than would have been ideal. We believe that this was possibly due to staff feeling fearful of completing the survey or skeptical of the anonymous nature of the data collection. While it is difficult to judge how this may have affected our results, it seems likely that this phenomenon would either be nondifferential or bias our findings to be an underestimation of the impact of TIPS, as nonparticipants would most likely hold more negative perceptions than those who chose to participate.

While research has focused on how interdisciplinary teams function in relation to clinical outcomes, additional work is needed to understand how to translate quality improvement efforts to sustained culture change for patient safety.[16] As health care reform is implemented over the coming decade with pay for performance and encouragement of collaboration across the continuum of care, it will be critical to have effective SNF teams partnering with hospitals, primary care medical homes, and home care agencies to manage transitions in care and avoid unnecessary hospitalizations. The results of this study suggest that a program such as the one described in this article could contribute to improved patient care by focusing the attention of SNF staff on improving quality and patient safety.

REFERENCES

1. Ouslander JG, Lamb G, Perloe M, et al. Potentially avoidable hospitalizations of nursing home residents: frequency, causes, and costs [see editorial comments by Drs. Jean F. Wyman and William R. Hazzard, pp 760-761]. *J Am Geriatr Soc.* 2010;58(4): 627-635.
2. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med.* 2009;360(14): 1418-1428.
3. Mor V, Intrator O, Feng Z, Grabowski DC. The revolving door of rehospitalization from skilled nursing facilities. *Health Aff.* 2010;29(1): 57-64.
4. Intrator O, Zinn J, Mor V. Nursing home characteristics and potentially preventable hospitalizations of long-stay residents. *J Am Geriatr Soc.* 2004;52(10):1730-1736.
5. Boockvar K, Fishman E, Kyriacou CK, Monias A, Gavi S, Cortes T. Adverse events due to discontinuations in drug use and dose changes in patients transferred between acute and long-term care facilities. *Arch Intern Med.* 2004;164(5):545-550.
6. Coleman EA, Smith JD, Frank JC, Min SJ, Parry C, Kramer AM. Preparing patients and caregivers to participate in care delivered across settings: the Care Transitions Intervention. *J Am Geriatr Soc.* 2004;52(11):1817-1825.

7. Naylor MD, Bowles KH, McCauley KM, et al. High-value transitional care: translation of research into practice [published ahead of print March 16, 2011]. *J Eval Clin Pract. doi:10.1111/j.1365-2753.2011.01659.x*
8. Jack BW, Chetty VK, Anthony D, et al. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. *Ann Intern Med.* 2009;150(3):178-187.
9. Harris R, Wilson-Barnett J, Griffiths P. Effectiveness of nursing-led inpatient care for patients with post-acute health care needs: secondary data analysis from a programme of randomized controlled trials. *J Eval Clin Pract*. 2007;13(2):198-205.
10. O'Reilly J, Lowson K, Green J, Young JB, Forster A. Post-acute care for older people in community hospitals—a cost-effectiveness analysis within a multi-centre randomised controlled trial. *Age Ageing*. 2008;37(5):513-520.
11. Kane RL, Keckhafer G, Flood S, Bershadsky B, Siaday MS. The effect of Evercare on hospital use. *J Am Geriatr Soc*. 2003;51(10):1427-1434.
12. Berkowitz RE, Jones RN, Rieder R, et al. Improving disposition outcomes for patients in a geriatric skilled nursing facility. *J Am Geriatr Soc*. 2011;59(6):1130-1136.
13. Bonner AF, Castle NG, Men A, Handler SM. Certified nursing assistants' perceptions of nursing home patient safety culture: is there a relationship to clinical outcomes? *J Am Med Dir Assoc*. 2009;10(1):11-20.
14. Nursing Home Survey on Patient Safety Culture. http://www.ahrq.gov/qual/patientsafetyculture/nhsurvindex.htm. Accessed May 20, 2011.
15. Boockvar KS, Burack OR. Organizational relationships between nursing homes and hospitals and quality of care during hospital-nursing home patient transfers. *J Am Geriatr Soc*. 2007;55(7):1078-1084.
16. Zwarenstein M, Reeves S, Russell A, et al. Structuring communication relationships for interprofessional teamwork (SCRIPT): a cluster randomized controlled trial. *Trials*. 2007;8:23.

Journal of Health Communication, 17:23–29, 2012

ISSN: 1081-0730 print/1087-0415 online
DOI: 10.1080/10810730.2012.712626

Routledge
Taylor & Francis Group

Commentary

The Role of Information Technology in Health Literacy Research

TIMOTHY W. BICKMORE

Northeastern University College of Computer and Information Science, Boston, Massachusetts, USA

MICHAEL K. PAASCHE-ORLOW

Section of General Internal Medicine, Boston Medical Center/Boston University School of Medicine, Boston, Massachusetts, USA

Without concerted effort, the current explosion in health information technology will further widen the digital health divide for individuals with inadequate health literacy. However, with focused investment of time and energy, technology has the potential for reducing disparities through intelligent, usable, and accessible systems that tailor information, advice, counseling, and behavioral support to an individual's need at a given time and place.

The development of information technology (IT) for patient and consumer health applications has been exploding in the past decade, with thousands of websites, hundreds of mobile applications, and dozens of special-purpose devices targeted at the health care market. However, this growth in technology is actually likely to increase health disparities for those with limited health, computer, and reading literacy, unless effort is devoted to the development of IT specifically designed for these disadvantaged groups, and issues of technology access are addressed. The design of technology for low-literacy individuals needs to be primarily focused on improved interface design, but all aspects of usability of health information systems must be addressed, including which health messages are delivered at a given time. There is now ample evidence that individuals with low health literacy have difficulty accessing and using state-of-the-art digital health communication media, such as websites (Jensen, King, Davis, & Guntzviller, 2010; Neter & Brainin, 2012; Neuhauser & Kreps, 2008; Sarkar et al., 2010). Barriers to health IT access, including cost, high-speed Internet access, access to consumer and patient medical devices, and rural-urban and regional variability in access, are often ignored by researchers developing new health technologies, but their resolution is crucial to ensure the ultimate success of their efforts.

Address correspondence to Timothy W. Bickmore, Northeastern University, Computer Science, WVH202, 360 Huntington Avenue, Boston, MA 02115, USA. E-mail: bickmore@ccs.neu.edu

If these issues of usability and access are addressed, IT has the potential to reduce health literacy-based disparities through the use of intelligent but easy-to-use systems that can inform, teach, persuade, and counsel, using messages that are perfectly tailored for an individual to understand and act upon in a given situation. There is now a growing body of research into the development and evaluation of computational artifacts designed specifically for populations with low health literacy that demonstrates the potential of IT to level the playing field by providing information at the place and time it is needed and in a format that is more understandable and actionable than traditional print media. Several of these artifacts have even been shown to be preferred over oral communication with health providers because they have the potential to provide a low-pressure context in which patients—especially those with low health literacy—do not feel talked down to and are open to asking questions (Bickmore, Pfeifer, & Jack, 2009).

In the rest of this article, we briefly survey the state of the art in using technology to help individuals with inadequate health literacy: automated assessment of individual health literacy level and document readability, and automated interventions for providing health information and/or counseling to individuals with inadequate health literacy. We close with some thoughts on future opportunities for the use of technology to assist individuals with low health literacy.

Information Technology in Health Literacy Assessment of Individuals

Several researchers have developed IT systems for automatically assessing the health literacy of individuals. Lobach and colleagues (Lobach, Arbanas, Mishra, Campbell, & Wildemuth, 2004; Lobach, Hasselblad, & Wildemuth, 2003) developed a computer-administered questionnaire to automatically assess users' reading and computer literacy levels, then used this assessment to automatically tailor the user interface to these factors. Yost and colleagues (2009) also developed a talking touchscreen system that assessed health literacy in English and Spanish, with the objective of completing assessments more efficiently and more accurately than the standard Test of Functional Health Literacy in Adults. To date, these efforts have been designed as research tools. In the near future, computer-adaptive testing will shorten the duration of these assessments and make them even easier to use. This could potentially lead to clinical use if and when a given patient's health literacy status needs to be identified in order to trigger specific interventions, although this has not yet occurred as of the time of writing.

Information Technology in Readability Assessment

Other researchers have developed automated approaches to assessing the readability levels of printed documents and websites. Although the Flesch-Kincaid readability calculator is the most well known given that it is a feature in Microsoft products, Dale-Chall, Gunning-Fog, SMOG, and other algorithms for automatically assessing the readability of documents are available in most modern word processors (Ley & Florio, 1996). These standard measures are easy to compute and use but are not reliable, as demonstrated by several recent studies (e.g., Feng, Jansche, Huenerfauth, & Elhadad, 2010). To date, they have been good tools to identify blocks of text that are likely to be inadequate. However, a transformation may be taking place as more sophisticated algorithms, on the basis of techniques from computational

linguistics have yielded more accurate results (Feng et al., 2010). Preliminary research has also been conducted into automatically mapping lay ("consumer") health care terms to standardized medical nomenclature (Zeng & Tse, 2006). This work lays the groundwork for systems that can eventually be developed to automatically translate between technical language and language that people can better understand.

Information Technology–Based Interventions

Several computerized health interventions have now been developed and evaluated specifically for individuals with low health literacy, such as multimedia touch screen computers that administer standard diagnostic questionnaires (Bryant et al., 2009; Wofford, Currin, Michielutte, & Wofford, 2001), and health education systems that are designed to accommodate a wide range of health, computer, and reading literacy levels (Gerber et al., 2005; Kandula et al., 2009). These systems commonly provide information using audio, video or graphical displays along with text to minimize the literacy level required of users, and often use touch screen input so that users do not need to have mouse and keyboard proficiency (Gerber et al., 2005; Kandula et al., 2009). For example, the Multimedia Diabetes Education Program uses images, 2D animation, and spoken audio (narrator) along with text. In a study of 190 patients (21% with low health literacy), it was found that all patients had significant increases in diabetes knowledge after interacting with the program, although the program did not fully compensate for low health literacy.

Interactive voice response (or phone-based) systems represent another modality that may be particularly well-suited to deliver interventions to low-literacy individuals because these systems (a) deliver information in speech instead of text and (b) use the ubiquitous communication channel of the telephone so that participants neither need computer literacy nor computer access (Piette, 2000). Schillinger and colleagues (2008) compared a multilingual interactive voice response system for diabetes self-management support with monthly group medical visits, and demonstrated that the interactive voice response system resulted in greater engagement, especially among participants with limited literacy. Shea and colleagues (2008) performed a direct comparison of a health satisfaction survey (Consumer Assessment of Healthcare Providers and Systems) delivered either in its original 63-item text format, the text format augmented with images, or a bilingual interactive voice response–based implementation, and demonstrated that significantly more individuals with inadequate health literacy completed the survey, and completed it is less time and with fewer errors, when delivered through interactive voice response compared with the other two modes.

Our own research in this area involves the use of embodied conversational agents that appear on the computer as animated characters that simulate the conversational nonverbal behavior that accompanies human face-to-face dialogue, such as hand gestures, facial displays, and gaze (Figure 1). We have found that patients learn more with embodied conversational agents compared to other modalities (e.g., human instructors), regardless of health literacy, and that individuals with inadequate health literacy generally report higher levels of satisfaction with embodied conversational agents and ask more questions with embodied conversational agents compared to individuals with adequate health literacy (Bickmore and Paasche-Orlow, 2011). However, as with the Multimedia Diabetes Education Program, embodied conversational agents did not fully compensate for low health literacy (Bickmore, Pfeifer, & Paasche-Orlow, 2009; Bickmore et al., 2010).

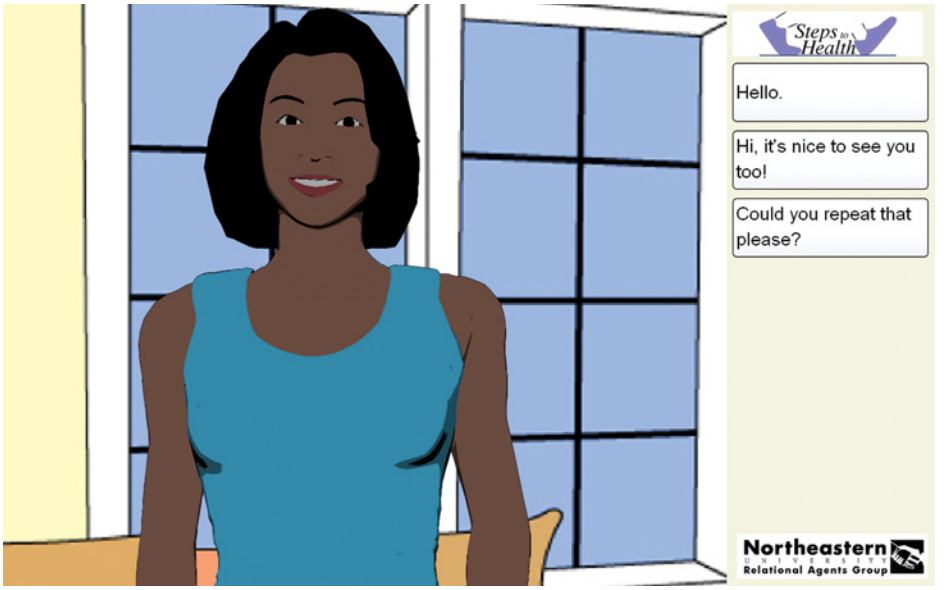


Figure 1. Example of embodied conversational agent used in walking promotion intervention for older adults (Bickmore et al., 2010). (Color figure available online.)

Future Directions for Information Technology in Health Literacy

There are a great many promising directions of research into how technology can improve the health of individuals with inadequate health literacy. As described earlier, most of the work to date in this area has been on the development of technologies designed to improve an individual's knowledge about specific areas of health and self-care, such as diabetes. Even in this area, there is a tremendous opportunity for new research into improved pedagogical methods that can be borrowed from the field of Intelligent Tutoring Systems and applied in specific health domains (Nkambou, Bourdeau, & Mizoguchi, 2010). In addition to improving knowledge, intelligent systems can be used to improve skill deficits in health care (e.g., inhaler technique) as well as underlying deficits in basic literacy skills. For example, automated reading tutors that listen to learners read out loud and provide coaching (Mostow et al., 2003) and math tutors that provide adaptive, personalized instruction to improve numeracy skills have both been developed and tested (Melis & Siekmann, 2004). IT systems can also play a role in providing access to health care and other services. For example, intelligent online directories have been designed to help connect urban, low income populations with community services including health care, housing, food security, and income security services (Fleegler, Lieu, Wise, & Muret-Wagstaff, 2007). Extending this kind of matchmaking service to add educational support for people who are struggling with their care is an important research direction.

Intelligent counseling systems can go beyond the provision of knowledge and literacy skills to impact a wide range of other psychological constructs associated with health literacy. Just as a good human counselor can work with an individual to improve self-efficacy and motivation to change a health behavior, automated

counseling agents can use techniques from health behavior change to improve health behaviors over time, and the messages and techniques used could be tailored to the individual's literacy (Bickmore & Giorgino, 2006; Bickmore, Schulman, & Sidner, 2011). Counseling agents could also be used to change norms by modeling or demonstrating ideal health behavior (e.g., medication adherence, walking, weight loss), and help prevent the discounting of future effects of current health decisions by reminding individuals of the long-term ramifications of their actions. Agents could also help address the fatalism about health outcomes many individuals with inadequate health literacy embrace by making individuals aware of the choices they have in their care and changing their expectations about the future.

Information systems could be used to significantly improve patient–provider communication for individuals with low health literacy. Interventions to promote patient activation are particularly important for low-literacy patients because they tend to ask fewer questions in comparison with individuals with adequate literacy (Katz, Jacobson, Veledar, & Kripalani, 2007; Schillinger et al., 2003). Mobile systems that maintain lists of questions and issues for patients and promote their resolution during medical encounters represent one of many ways that IT systems could become automated advocates for those who lack family or friends who can perform this role. Automated coaches could also give patients practice interacting with virtual health providers to build self-efficacy for interactions with the actual health care system. Automated systems could also be used to help facilitate and mediate patient–provider interactions by, for example, negotiating the agenda of problems to address with providers and by facilitating shared decision making. Automated systems could also be used to change the behavior and attitudes of providers themselves when interacting with low-literacy patients by, for example, prompting them to evaluate comprehension and other best practice interventions.

Last, IT systems can be used to significantly enhance the social support networks for individuals with low health literacy. Existing websites such as PatientsLikeMe.com and the many condition-specific support websites are wonderful resources for those who are connected into the digital world, but these resources are largely out of reach for those with low computer literacy and limited resources. In addition, the members of these online communities tend be well educated and highly health literate, which may represent a barrier for low-literacy individuals to fully participate. Parallel resources need to be established for those who lack this digital and health fluency and lack access to IT systems to connect them both to expert health advice and peers who can provide emotional, informational, and instrumental support.

Conclusion

Imagine a future in which an individual can carry a mobile device (perhaps their smartphone) that not only provides health information, but senses and interprets the environment and makes proactive recommendations about actions that might affect their health. The device could also translate health-related text into actionable messages, tailored not only to the individual's health literacy level, but their educational and cultural background, values, and even emotional state, providing information and recommendations at the time and location, and in the format, that will have the most effect. Such a digital health conscience would need a detailed cognitive model of its wearer, going beyond canonical models of health literacy and culture—which can simply reinforce stereotypes—to a detailed representation of a

user's unique desires, beliefs, and behavior patterns, and how these have changed over time. Elements of such systems already exist, and limited versions are certainly in our near future. However, the market for health IT products will not take care of our most vulnerable populations on its own; we need continued, focused, evidence-based research to continue developing technologies that specifically address disparities for individuals with low health literacy in order to realize this potential.

References

Bickmore, T., & Giorgino, T. (2006). Health dialog systems for patients and consumers. *Journal of Biomedical Informatics*, *39*, 556–571.

Bickmore, T., & Paasche-Orlow, M. (2011). *Health Literacy and Therapeutic Alliance with Relational Computer Agents*. Paper presented at the Health Literacy Annual Research Conference, Chicago, Illinois.

Bickmore, T., Pfeifer, L., Byron, D., Forsythe, S., Henault, L., Jack, B., … Paasche-Orlow, M. (2010). Usability of conversational agents by patients with inadequate health literacy: Evidence from two clinical trials. *Journal of Health Communication*, *15*, 197–210.

Bickmore, T., Pfeifer, L., & Jack, B. W. (2009). Taking the time to care: Empowering low health literacy hospital patients with virtual nurse agents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 1265–1274), Boston, MA. New York, NY: Association for Computing Machinery.

Bickmore, T., Pfeifer, L., & Paasche-Orlow, M. (2009). Using computer agents to explain medical documents to patients with low health literacy. *Patient Education and Counseling*, *75*, 315–320.

Bickmore, T., Schulman, D., & Sidner, C. (2011). A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of Biomedical Informatics*, *44*, 183–197.

Bryant, M. D., Schoenberg, E. D., Johnson, T. V., Goodman, M., Owen-Smith, A., & Master, V. A. (2009). Multimedia version of a standard medical questionnaire improves patient understanding across all literacy levels. *Journal of Urology*, *182*, 1120–1125.

Feng, L., Jansche, M., Huenerfauth, M., & Elhadad, N. (2010). A comparison of features for automatic readability assessment. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING)* (pp. 276–284). Stroudsburg, PA: Association for Computational Linguistics.

Fleegler, E. W., Lieu, T. A., Wise, P. H., & Muret-Wagstaff, S. (2007). Families' health-related social problems and missed referral opportunities. *Pediatrics, 119,* e1332–e1341.

Gerber, B. S., Brodsky, I. G., Lawless, K. A., Smolen, L. I., Arozullah, A. M., Smith, E. V., … Eiser, A. R. (2005). Implementation and evaluation of a low-literacy diabetes education computer multimedia application. *Diabetes Care*, *28*, 1574–1580.

Jensen, J., King, A., Davis, L., & Guntzviller, L. (2010). Utilization of internet technology by low-income adults: The role of health literacy, health numeracy, and computer assistance. *Journal of Aging and Health*, *22*, 804–826.

Kandula, N. R., Nsiah-Kumi, P. A., Makoul, G., Sager, J., Zei, C., Glass, S., … Baker, D. (2009). The relationship between health literacy and knowledge improvement after a multimedia type 2 diabetes education program. *Patient Education and Counseling*, *75*, 321–327.

Katz, M. G., Jacobson, T. A., Veledar, E., & Kripalani, S. (2007). Patient literacy and question-asking behavior during the medical encounter: A mixed-methods analysis. *Journal of General Internal Medicine*, *22*, 782–786.

Ley, P., & Florio, T. (1996). The use of readability formulas in health care. *Psychology, Health & Medicine*, *1*, 7–28.

Lobach, D., Arbanas, J., Mishra, D., Campbell, M., & Wildemuth, B. (2004). Adapting the human-computer interface for reading literacy and computer skill to facilitate collection of information directly from patients. *Studies in Health Technology and Informatics*, *107*, 1142–1146.

Lobach, D., Hasselblad, V., & Wildemuth, B. (2003). Evaluation of a tool to categorize patients by reading literacy and computer skill to facilitate the computer-administered patient interview. In *Proceedings of the Annual Symposium of the American Medical Informatics Association (AMIA)* (pp. 391–395), Washington, DC. Bethesda, MD: AMIA.

Melis, E., & Siekmann, J. (2004). ActiveMath: An intelligent tutoring system for mathematics. In L. Rutkowski, J. Siekemann, R. Tadeusiewicz, & L. A. Zadeh (Eds.), *Proceedings of the Seventh International Conference "Artificial Intelligence and Soft Computing" (ICAISC)* (pp. 91–101). Springer-Verlag.

Mostow, J., Aist, G., Burkhead, P., Corbett, A., Cuneo, A., Ettelman, S., … Tobin, B. (2003). Evaluation of an automated reading tutor that listens: Comparison to human tutoring and classroom instruction. *Journal of Educational Computing Research*, *29*(1), 61–117.

Neter, E., & Brainin, E. (2012). eHealth literacy: Extending the digital divide to the realm of health information. *Journal of Medical Internet Research, 14*, e19. Retrieved from http://www.jmir.org/2012/1/e19/

Neuhauser, L., & Kreps, G. (2008). Online cancer communication: Meeting the literacy, cultural and linguistic needs of diverse audiences. *Patient Education and Counseling*, *71*, 365–377.

Nkambou, R., Bourdeau, J., & Mizoguchi, R. (2010). *Advances in Intelligent Tutoring Systems.* New York, NY: Springer.

Piette, J. (2000). Interactive voice response systems in the diagnosis and management of chronic disease. *American Journal of Managed Care*, *6*, 817–827.

Sarkar, U., Karter, A., Liu, J., et al. (2010). The literacy divide: Health literacy and the use of an Internet-based patient portal in an integrated health system—Results from the Diabetes Study of Northern California (DISTANCE). *Journal of Health Communication*, *15*, 183–196.

Schillinger, D., Hammer, H., Wang, F., et al. (2008). Seeing in 3-D: Examining the reach of diabetes self-management support strategies in a public health care system. *Health Education & Behavior*, *35*, 664–682.

Schillinger, D., Piette, J., Grumbach, K., Wang, F., Wilson, C., & Daher, C. (2003). Closing the loop: Physician communication with diabetic patients who have low health literacy. *Archives of Internal Medicine*, *163*, 83–90.

Shea, J., Guerra, C., Weiner, J., Aguirre, A., Ravenell, K., & Asch, D. (2008). Adapting a patient satisfaction instrument for low literate and Spanish-speaking populations: Comparison of three formats. *Patient Education and Counseling*, *73*, 132–140.

Wofford, J., Currin, D., Michielutte, R., & Wofford, M. (2001). The multimedia computer for low-literacy patient education: A pilot project of cancer risk perceptions. *Medscape General Medicine, 3,* 23. Retrieved from http://www.medscape.com/viewarticle/408129

Yost, K. J., Webster, K., Baker, D. W., Choi, S. W., Bode, R. K., & Hahn, E. A. (2009). Bilingual health literacy assessment using the Talking Touchscreen/la Pantalla Parlanchina: Development and pilot testing. *Patient Education and Counseling*, *75*, 295–301.

Zeng, Q., & Tse, T. (2006). Exploring and developing consumer health vocabularies. *JAMA*, *13*, 24–29.

Blood and Cheng *BMC Medical Research Methodology* 2012, **12**:5
http://www.biomedcentral.com/1471-2288/12/5

BMC
Medical Research Methodology

RESEARCH ARTICLE Open Access

Non-linear mixed models in the analysis of mediated longitudinal data with binary outcomes

Emily A Blood[1,2*] and Debbie M Cheng[1]

Abstract

Background: Structural equation models (SEMs) provide a general framework for analyzing mediated longitudinal data. However when interest is in the total effect (i.e. direct plus indirect) of a predictor on the binary outcome, alternative statistical techniques such as non-linear mixed models (NLMM) may be preferable, particularly if specific causal pathways are not hypothesized or specialized SEM software is not readily available. The purpose of this paper is to evaluate the performance of the NLMM in a setting where the SEM is presumed optimal.

Methods: We performed a simulation study to assess the performance of NLMMs relative to SEMs with respect to bias, coverage probability, and power in the analysis of mediated binary longitudinal outcomes. Both logistic and probit models were evaluated. Models were also applied to data from a longitudinal study assessing the impact of alcohol consumption on HIV disease progression.

Results: For the logistic model, the NLMM adequately estimated the total effect of a repeated predictor on the repeated binary outcome and were similar to the SEM across a variety of scenarios evaluating sample size, effect size, and distributions of direct vs. indirect effects. For the probit model, the NLMM adequately estimated the total effect of the repeated predictor, however, the probit SEM overestimated effects.

Conclusions: Both logistic and probit NLMMs performed well relative to corresponding SEMs with respect to bias, coverage probability and power. In addition, in the probit setting, the NLMM may produce better estimates of the total effect than the probit SEM, which appeared to overestimate effects.

Background

SEMs are a general modeling framework often used in the social sciences to analyze complex relationships between variables, such as mediated relationships between variables. A mediator is a variable in the causal pathway between a predictor and the outcome of interest. SEMs are becoming more common in the clinical research setting and can be used to model hypothesized causal pathways between variables of interest. Extensions of SEMs have been developed to allow for more general types of dependent variables, including binary outcomes [1]. Common statistical techniques for non-mediated longitudinal binary data include non-linear mixed models (NLMM) [2] and generalized estimating equations (GEE) [3]. When interest is primarily in the total effect of a predictor on an outcome, even if mediation may be present, these commonly used techniques may be preferred over SEMs as they specify straightforward predictor-outcome variable relationships and do not require specialized software, as the SEM often does. It is therefore of interest to determine, in a setting conducive to using SEMs, whether a method such as NLMMs adequately models the total effect of a predictor on binary outcomes without directly modeling mediation. We focus on NLMM rather than GEE in this paper as it is more similar to the non-linear SEMs for longitudinal data available in SEM software-both are conditional rather than marginal models.

Comparisons have been made between SEM and other statistical models in different contexts [4-13]. Mixed effect models have been evaluated against SEMs with continuous data [14,15], and found to adequately model mediated predictor-outcome relationships. MacKinnon et al. [16] examined the calculation of mediated effects in cross-sectional binary data with non-SEM techniques using two different methods (difference of coefficients

* Correspondence: emily.blood@childrens.harvard.edu
[1]Department of Biostatistics Boston University School of Public Health 801 Massachusetts Avenue 3rd Floor Boston, MA 02118 USA
Full list of author information is available at the end of the article

and product of coefficients). While, Palta and Lin [17] compared structural equation models to various marginal models in longitudinal binary data without mediation. To our knowledge, evaluation of NLMMs relative to SEMs has not yet been performed in the context of mediated longitudinal binary data.

Linear and non-linear mixed models differ both in terms of the distributional assumptions and the estimation techniques used for inference. In addition, the parameter estimates in non-linear mixed models using a logit or probit link are inherently scaled to the predictors (and mediators) included in the model. Therefore, comparisons of parameter estimates between NLMMs with different sets of predictors must first be re-scaled in order to make them comparable [16].

In this paper, we evaluate the performance of NLMMs relative to SEMs for the modeling of mediated, binary longitudinal data in a setting where the SEM is presumed to be optimal. The purpose is to assess whether there is an impact of direct modeling of causal pathways in terms of bias, power, and coverage probability when the goal is to determine the total effect of the main independent variable. A simulation study is performed to assess these two classes of models across a variety of settings. We also describe, in an appendix, two different approaches for rescaling estimates when analyzing real world data in order to allow direct comparisons between NLMMs and SEMs or to compute mediated effects via NLMMs only.

Methods

In the current study, we consider a longitudinal data setting with binary outcomes, a repeated binary predictor, a repeated continuous mediator, and a continuous covariate measured at baseline. An example of such a clinical setting would be a prospective cohort study evaluating the impact of heavy alcohol consumption on HIV disease progression, defined as low CD4 cell count (e.g. <350 cells/μL). Heavy alcohol consumption may influence progression of HIV, while also influencing adherence to anti-retroviral therapy (ART). Level of adherence to ART is also a predictor of HIV disease progression. In this setting there is a repeated binary independent variable of primary interest, heavy alcohol consumption (z_j), and a longitudinal binary outcome, low CD4 cell count (Y_j) signifying HIV progression. In addition, ART adherence (M_j), a continuous mediating variable, is measured repeatedly, and age (w) is a continuous covariate assessed at baseline. ART adherence is said to be a mediator because the primary independent variable, heavy alcohol use, may affect CD4 count directly as well as indirectly through ART adherence. We arbitrarily assume six timepoints at which the predictor, outcome and mediator are measured. Time is represented by t_j with j = 1, 2,..., 6. In this setting, we considered measurement times to be equally spaced and the same for all individuals. We generated data with a mediated non-linear relationship between the predictor (heavy alcohol consumption) and outcome (low CD4 cell count), i.e. we allowed the mediator (ART adherence) to be directly affected by the predictor and the outcome to be directly affected by both the predictor and mediator. Both the probit and logit links were assessed. We also describe the application of these models to data from a prospective cohort study evaluating the impact of heavy alcohol use on HIV disease progression.

As described by Fitzmaurice, Laird, and Ware [18] and others, binary outcome models can be described equivalently in two ways. The first approach would be to define a linear function of an underlying latent continuous variable (Y^*) that when dichotomized represents the observed binary outcome (Y). For example, we could define a continuous latent (unobserved) variable Y^* such that observed Y is 1 if $Y^* > 0$ and 0 otherwise and write:

$$Y^* = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

A second approach would be to define a non-linear model of the probability of a binary response. If we consider $\epsilon \sim N(0,1)$, the model in Equation 1 defines a univariate probit model that can be equivalently represented using the following non-linear link format:

$$P(Y = 1) = \Phi(\beta_0 + \beta_1 x).$$

Likewise, if we could consider the errors in Equation 1 to have standard logistic distribution (mean of 0 and variance of $\frac{\pi^2}{3}$) the model defines a univariate logistic model that can be represented as:

$$P(Y = 1) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}.$$

In more complex situations, such as the longitudinal data we are studying, the same equivalence between model descriptions exist and we use both model formulations for the NLMMs and SEMs that follow. The convention for binary or categorical outcomes in SEMs has been to describe binary regression models with the latent variable format while the NLMMs are often defined using the non-linear link format.

SEM

To evaluate the performance of NLMMs in a setting conducive to the use of SEMs, we generated mediated longitudinal binary outcomes using a non-linear SEM. We then fit the data with a NLMM as well as the non-

linear SEM to evaluate the performance of the NLMM relative to the SEM. The non-linear SEM used to generate the data and subsequently fit to the generated data is described below.

Following the notation from above, x_j is the independent variable of primary interest, M_j is the continuous mediating variable, w is a continuous time-invariant covariate, and t_j represents time-point. Using the latent variable notation, we define a continuous unobserved outcome Y_{ij}^* that takes a value of 1 only if $Y_{ij}^* > 0$ for j = 1 to 6. This model can be expressed as follows (dropping the subject index i for simplicity), where:

Measurement model

$$Y_j^* = U_1 + t_j U_2 + \lambda M_j + \kappa z_j + \varepsilon_j \quad (2)$$

Just as in the simpler models above, if we assume $\epsilon_j \sim$ N(0,1) this defines a probit model and if we assume $\epsilon_j \sim$ Logistic $(0, \frac{\pi^2}{3})$ this defines a logit model.

Structural model

$$U_1 = \alpha_1 + \gamma_2 w + \varsigma_1 \quad (3)$$

$$U_2 = \alpha_2 + \varsigma_2 \quad (4)$$

For j = 1 to 6,

$$M_j = \alpha_3 + \gamma_1 z_j + \varsigma_{2+j}, \quad (5)$$

where U_{i1} represents a latent intercept, U_{i2} represents a latent slope, z_{ij} represents the repeated binary predictor and M_{ij} represents the repeated continuous mediator. The errors in the structural model are normally distributed with cov(ζ_1, ζ_2) = Ψ, cov($\zeta_3 : \zeta_8$) = Θ and $\zeta_{(2+j)} \sim$ N(0, θ). This model can be represented in a path diagram (Figure 1), a visual display of the interrelationships between variables typically presented along with SEMs.

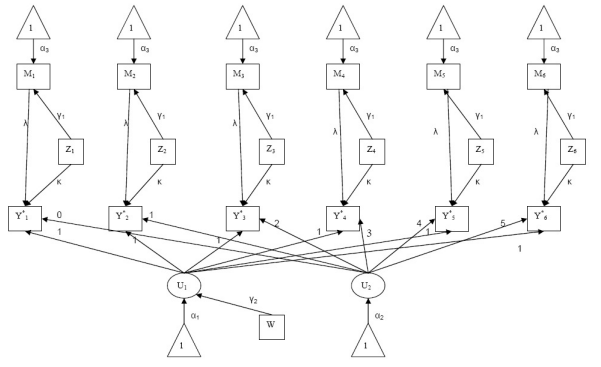


Figure 1 Path diagram. The non-linear structural equation model defined in Equations 2 to 5.

The parameters of the SEM defined in Equations 2 - 5 include: λ, which represents the effect of the repeated mediator on the repeated outcome; γ_1, which represents the effect the repeated primary independent variables on the repeated mediator; γ_2, which represents the effect of the continuous covariate on the repeated outcomes; and κ, which represents the effect of the repeated independent variable on the repeated outcome.

In this simulation study we focused on the total effect of the repeated binary predictor and the repeated binary outcome, which is represented by $\lambda\gamma_1 + \kappa$. The interpretation of the parameters of this model is subject-specific since it represents the effect of a predictor on the outcome when the individual intercept, individual slope and mediator value are held constant.

When the structural model Equations (3-5) are substituted into the measurement model Equation 2, the full model can be rewritten as:

$$Y_j^* = \omega_j + \varsigma_1 + t_j \varsigma_2 + \lambda \varsigma_{2+j} + \varepsilon_j$$

where $\omega_j = (\alpha_1 + \lambda\ \alpha_3) + \gamma_2 w + \alpha_2\ t_j + (\kappa + \lambda\gamma_1)\ z_j$. The following presents the non-linear link formats for the probit and logit SEMs where the structural equations have been substituted into the measurement equation (the subject index i has again been dropped for simplicity):

Probit SEM

$$\mathrm{E}(\mathrm{Y}_1|\varsigma_1, \varsigma_2, \varsigma_{2+j}) = \Phi(\omega_j + \varsigma_1 + t_j \varsigma_2 + \lambda \varsigma_{2+j}) \quad (6)$$

Logit SEM

$$E(Y_j|\varsigma_1, \varsigma_2, \varsigma_{2+j}) = \frac{\exp(\omega_j + \varsigma_1 + \mathrm{t_j}\varsigma_2 + \lambda\varsigma_{2+j})}{1 + \exp(\omega_j + \varsigma_1 + \mathrm{t_j}\varsigma_2 + \lambda\varsigma_{2+j})} \quad (7)$$

To fit these models, Mplus uses maximum likelihood estimation when a logit link is used and weighted least squares estimation with a robust estimation of standard errors (WLSMV) when a probit link is used [19].

Non-linear mixed effects model

The following NLMM was evaluated in comparison to the SEM:

$$Y_j^* = \beta_0 + \beta_1 w + \beta_2 t_j + \beta_3 z_j + b_1 + b_2 t_j + e_j. \quad (8)$$

where b_1 is a random individual intercept and b_2 is a random individual slope. Since the objective is to evaluate the total effect of the main independent variable, the mediator is excluded from this model [20]. The regression coefficient associated with the primary predictor

(β_3) therefore represents its total (i.e. direct plus indirect) effect on the outcome [14].

Probit NLMM

The probit model assumes that $e_j \sim N(0,1)$ and can be written as:

$$E(Y|b_1, b_2) = \Phi(\nu_j + b_1 + b_2 t_j) \quad (9)$$

where $\nu_j = \beta_0 + \beta_1 w + \beta_2 t_j + \beta_3 z_j$.

Logit NLMM

the logit model assumes that $e_j \sim$ Logistic $(0, \sqrt{\frac{\pi^2}{3}})$ and can be written as:

$$E(Y|b_1, b_2) = \frac{\exp(\nu_j + b_1 + b_2 t_j}{1 + \exp(\nu_j + b_1 + b_2 t_j)} \quad (10)$$

These models can be fit with SAS PROC NLMIXED which estimates parameters via maximum likelihood [21]. We note that the regression coefficients of the NLMM are interpreted conditional on the random individual intercept and random individual slope, but marginal on the residual error of the mediator (since the mediator is not included in the model).

Comparing NLMMs to SEMs

As noted previously, the SEM and NLMM condition differently on the mediating variable. Specifically, the SEM conditions on the random intercept and slope as well as on the residual variance of the mediating variable, while the NLMM conditions only on the random intercept and random individual slope. Thus estimates from the two types of models are not directly incomparable. Instead, to compare parameters from the NLMM to that of the SEM, we must first re-scale the regression coefficient from the NLMM so that it represents the effect of the primary predictor variable z_j conditional on the mediator. To determine the scaling factor, we rewrite the SEM (for both the probit and logistic models) conditional only on the variance of the random intercept and slope to mimic the conditioning in the NLMM.

Comparing probit models

For the probit SEM, we generated the data according to the model described in Equations 6 and 7. Conditioning only on the variances of the random intercept and slope (ζ_1 and ζ_2), but not on the variance of the mediator (ζ_{2+j}), it can be shown that:

$$\begin{aligned} E(Y|\varsigma_1, \varsigma_2) &= P(Y^* > 0|\varsigma_1, \varsigma_2) \\ &= P(\varepsilon_j + \lambda\varsigma_{2+j} > -(\omega_j + \varsigma_1 + t_j\varsigma_2)|\varsigma_1, \varsigma_2). \end{aligned} \quad (11)$$

The sum of terms on the left-hand side of the inequality do not have a standard normal distribution since $\lambda\zeta_{2+j}$ is added to ϵ_j which itself has a standard normal distribution. In order to express the probability in Equation 11 using the standard normal cumulative probability function, we re-scale the terms on either side of the inequality by the standard deviation of $\epsilon_j + \lambda\zeta_{2+j}$ to create a standard normal random variable:

$$\begin{aligned} &= P\left(\frac{\varepsilon_j + \lambda\varsigma_{2+j}}{\sqrt{1+\lambda^2\theta}} > \frac{-(\omega_j + \varsigma_1 + t_j\varsigma_2)}{\sqrt{1+\lambda^2\theta}} \,|\varsigma_1, \varsigma_2\right) \\ &= \Phi\left(\frac{(\omega_j + \varsigma_1 + \varsigma_2 t_j}{\sqrt{1+\lambda^2\theta}}\right) \end{aligned}$$

Conditioning on only the variance of the random individual intercept and slope, all regression coefficients are divided by the factor $\sqrt{1+\lambda^2\theta}$. For example, the regression coefficient associated with z_j, which was $\kappa + \lambda\ \gamma_1$, is now $\frac{\kappa + \lambda\gamma_1}{\sqrt{1+\lambda^2\theta}}$. Thus, the model parameters from the SEM are scaled to the variance of $\epsilon_j + \lambda\zeta_{2+j}$ which is $1 + \lambda^2\ \theta$ and the model parameters from the NLMM are scaled to the variance of ϵ_j which is 1, resulting in a scaling factor of $\sqrt{\frac{1+\lambda^2\theta}{1}}$ Parameter estimates from the SEM and NLMM must be on the same scale before making direct comparisons. For example, the total effect of the main independent variable from the probit NLMM, β_3, which is also conditioned only on the random individual intercept and slope (Equation 9) should be multiplied by a factor of $\sqrt{1+\lambda^2\theta}$ before it is compared to the total effect from the probit SEM, $\kappa + \lambda\gamma_{71}$. Direct comparisons of parameter estimates from the NLMM to those from the SEM without first re-scaling would underestimate effects by a factor of $\sqrt{1+\lambda^2\theta}$. In the current study we present the conditional total effect estimates from the SEM and compare them to scaled and unscaled NLMM estimates. Note that in the analysis of real (i.e. non-simulated) data, true parameter values are unknown and therefore must be estimated. We describe in the appendix two approaches for rescaling estimates in practice to allow direct comparisons between NLMMs and SEMs or to compute mediated effects via NLMMs only.

Comparing logistic models

Unlike the probit model, when the logit SEM is conditioned on only the random intercept and slope, the true relationship between the predictor and the outcome no longer follows a logistic model. That is, the distribution of the terms on the left-hand side of Equation 11 in a logit SEM does not follow a logistic distribution since the sum of a normal random variable (ϵ_j) and logistic random variable (ζ_{2+j}) does not follow a logistic distribution. The result of this is that the scaled coefficients from the logit NLMM only approximate the mediated relationship described in a logit SEM. A similar situation occurs, for example, when comparing a non-linear

mixed model to a non-linear generalized estimating equation as noted by Fitzmaurice, Laird, and Ware [18].

The scale factor for the logit model is created in the same was as it was for the probit model. The regression coefficient representing the total effect of the main independent variable (β_3) from the logit NLMM, can be multiplied by the standard deviation of $\epsilon_j + \lambda\zeta_{2+j}$ and divided by the standard deviation of ϵ_j. The scaling factor for the logit model is therefore: $\left[\left(\frac{\pi^2}{3} + \lambda^2\theta\right)/\frac{\pi^2}{3}\right]^{\frac{1}{2}}$.

Simulation plan

Data generation and model fitting

Because the goal of this study was to evaluate the performance of NLMM relative to SEMs in the setting where the SEM is presumed to be optimal, the SEM framework was used to generate the mediated binary data for the simulation studies. Data were generated according to Equations 2-5. For the probit model, errors in Equation 2 were assumed to be independent standard normal random variates. For the logistic model, the errors in Equation 2 were assumed to be independent standard logistic random variates. Data generation was repeated to create 1000 datasets. NLMM were fit with SAS (Version 9.2) PROC NLMIXED and SEM were fit with Mplus (Version 5.2).

Simulated data scenarios

We evaluated the performance of NLMM against SEM across several scenarios by examining the following:

- Sample size: ranging from 100 to 1000. The range of sample sizes was chosen to evaluate sample sizes that achieved adequate power with a moderate effect size.
- Effect size: ranging from 0.2 to 0.5. The range of effect sizes represent small to moderate effect sizes as defined by Cohen [22].
- Distribution of effects: three cases were evaluated: equally distributed direct and indirect effects, primarily direct effects of the main independent variable, and primarily indirect effects of the main independent variable. The total effect sizes (0.3 for probit models and 0.4 for logit models) were chosen such that adequate power was obtained when direct and indirect effects were equally distributed.

Model performance was assessed based on the following: 1) Bias- the difference between the true parameter value and the mean observed parameter value divided by the true parameter value; 2) Coverage probability- the percentage of the 1000 95% confidence intervals that contained the true parameter value; 3) Power - the percentage of the 1000 datasets in which the null hypothesis that the total effect of the main independent variable is equal to zero was statistically significant.

Results

Logistic link

The results evaluating the conditional total effect from simulations evaluating sample size are displayed in Table 1. For the logistic model, a sample size of 700 was required to obtain adequate (83%) power to detect a moderate effect size (0.3) for both the SEM and NLMM. Across all sample sizes, the estimates of power for the scaled NLMM were comparable to the SEM. As expected, for the unscaled NLMM, effects were underestimated in all cases. For example with a sample size of 600, the bias of the SEM was -0.6% while the bias for the unscaled NLMM was -3.8% and the bias of the scaled NLMM was -1.6%. Once the scale factor was applied to the NLMM, however, the estimated bias decreased and the power and coverage probability estimates were similar to that of the SEM. The results for the unscaled NLMM are therefore not described in subsequent tables as they are not directly comparable to the SEM results.

The results of the simulations in which the effect size (Table 2) and effect distribution (Table 3) were varied also demonstrated that the appropriately scaled NLMM produced comparable estimates of the total effect of the exposure relative to the SEM. The scaled NLMM results were similar to the comparison SEM with low bias at all effect sizes (ranging from -2.0% to 0.4%) and effect distributions (ranging from -1.4% to -1.0%). As expected, the models showed increasing power with increasing effect sizes (e.g., from 37% power for an effect size of 0.2 to 98% power for an effect size of 0.5) and power was highest (91%) when the effect distribution was primarily direct.

Probit models

For the probit model, the SEM showed consistent positive bias (i.e., no simulation scenario with the probit model resulted in a negative bias for the probit SEM). For example, the estimated bias for a sample size of 100 was 51.5% and decreased to 7.5% with a sample size of 1000. In comparison, the scaled NLMM had bias ranging from 2.7% to -2.1%. Notably, the estimated power for the NLMM was consistently higher than that of the comparison SEM. In the effect size simulation scenarios, bias in the SEM appeared to increase with effect size (effect sizes of 0.2, 0.3, and 0.5 resulted in biases of 7.6%, 8.8%, and 11.7%, respectively). In contrast, the

Table 1 Impact of sample size. **Based on 1000 simulated datasets with moderate effect size (0.3) equally distributed between direct and indirect effects. Impact of sample size on model performance in evaluating the total effect of the repeated independent variable on the repeated outcome.**

Simulated Data	SEM			Unscaled NLMM			Scaled NLMM		
Sample Size	Bias (%)	Coverage (%)	Power (%)	Bias (%)	Coverage (%)	Power (%)	Bias (%)	Coverage (%)	Power (%)
Logit Link Results									
200	1.3	95	35	-2.3	95	34	-0.09	95	35
300	2.3	96	49	-0.8	96	47	1.4	96	48
400	-0.8	95	57	-4.1	95	56	-1.9	95	56
500	0.2	94	68	-2.8	95	68	-0.6	95	68
600	-0.6	94	77	-3.8	94	77	-1.6	94	77
700	0.1	94	83	-3.0	95	83	-0.8	94	84
Probit Link Results									
100	51.5	97	20	-4.7	94	31	2.1	94	30
200	20.4	96	42	-4.3	94	53	2.7	94	54
300	11.4	95	57	-7.1	94	72	-0.2	95	72
400	8.3	93	69	-7.9	94	80	-1.2	94	80
500	8.8	95	79	-6.7	93	88	0.2	94	88
600	7.5	94	87	-7.0	93	93	-0.2	94	94
1000	7.5	94	87	-8.8	92	99	-2.1	94	99

probit NLMM showed relatively small bias (-0.7% to 1.7%) for all effect sizes. For the probit model, the scaled NLMM generally performed better than the SEM, across a range of sample sizes, effect sizes, and effect distributions, with higher estimated power and lower bias.

Positive bias in probit models

We explored possible explanations of the unexpected positive bias observed for the probit SEM models. We repeated simulations by first eliminating mediation from the simulated scenarios to evaluate whether the positive bias in the SEM persisted in unmediated settings. In addition, we assessed different estimation methods, including WLSMV (implemented in MPlus software) and also maximum likelihood using iteratively reweighted least squares (implemented in Splus and SAS). We found that the mean regression coefficient estimates using WLSMV had larger bias in nearly all cases compared to estimates using maximum likelihood and iteratively reweighted least squares (Table 4). The positive bias for the WLSMV method was avoided only with a sample size of 5000. Our results suggest that the WLSMV, which is used by Mplus, may produce positively biased results. This bias was also present, however to a smaller degree, in the NLMM fit with maximum likelihood.

Table 2 Impact of effect size. Based on 1000 simulated datasets with sample size of 500 equally distributed between direct and indirect effects. Impact of effect size on model performance in evaluating the total effect of the repeated independent variable on the repeated outcome.

Simulated Data	SEM			Scaled NLMM		
Effect Size	Bias (%)	Coverage (%)	Power (%)	Bias (%)	Coverage (%)	Power (%)
Logit Link Results						
0.2	-1.1	96	38	-2.0	96	37
0.3	0.2	94	68	-0.6	95	68
0.4	-0.6	96	89	-1.3	95	89
0.5	1.3	95	99	0.4	95	98
Probit Link Results						
0.2	7.6	94	47	-0.7	95	58
0.3	8.8	95	79	0.2	94	88
0.4	11.1	93	95	1.7	93	98
0.5	11.7	95	>99	1.6	95	>99

Table 3 Impact of effect distribution. Based on 1000 simulated datasets with sample size of 500 and effect size of 0.4 for the logit link and effect size of 0.3 for the probit link. Impact of effect distribution on model performance in evaluating the total effect of the repeated independent variable on the repeated outcome.

Simulated Data	SEM			Scaled NLMM		
Effect Distribution	**Bias (%)**	**Coverage (%)**	**Power (%)**	**Bias (%)**	**Coverage (%)**	**Power (%)**
Logit Link Results						
Equal	-0.6	96	89	-1.3	95	89
Direct	-0.6	95	91	-1.0	95	91
Indirect	-0.2	95	90	-1.4	95	89
Probit Link Results						
Equal	8.8	95	79	0.2	94	88
Direct	5.9	96	80	-0.9	95	90
Indirect	8.6	95	76	0.09	94	88

Real data example: alcohol and HIV disease progression

To demonstrate the application of both the logit and probit NLMMs and SEMs evaluated in the simulation study, we analyzed data from a prospective cohort study evaluating the effect of alcohol use on HIV disease progression. Samet et al. have previously reported the analyses from this longitudinal cohort study [23]. The original analyses combined data from two cohorts (the HIV-ALC and HIV-LIVE cohorts), however, to illustrate the models evaluated in this paper, we have used data from the HIV-LIVE study only. For clarity of presentation, we limited the analyses to subjects who reported any ART use during follow-up, had complete data on the first four time-points (as Mplus and SAS have different methods for handling missing data in these models), and examined only the following key variables: heavy alcohol consumption (yes vs. no), the main independent variable; ART adherence (percentage of pills taken in the last three days), the mediator; age, a potential confounder; and low CD4 cell count (dichotomized at <350 cells/μL), the primary outcome. Each variable was assessed every six months for up to four years, however for the current example only the first four time-points were analyzed in order to maximize the number of subjects with complete data. The resulting dataset was composed of 98 individuals contributing 392 observations. The total effect of heavy alcohol consumption on low CD4 cell count was not significant for any of the SEMs or NLMMs fit to the data. For the logit SEM, the total effect estimate (SE) was 0.554(1.246) with an associated p-value of 0.66. The scaled result from the logit NLMM was similar: estimated total effect (SE) = 0.5107(0.701), p = 0.47. However for the probit link, the estimated total effects (SE) from the SEM and NLMM appeared to differ substantially (probit SEM: 6.287(52.661),p = 0.91; scaled probit NLMM: 0.303(0.391), p = 0.44). Thus consistent with the results from the simulation study, the logit SEM and NLMM produced similar estimates in the real data example, whereas the probit SEM produced estimated effects that appeared much larger in magnitude in comparison to the probit NLMM.

Table 4 Univariate Probit Model Results

Sample Size	Effect Size	WLSMV	Bias ML-IRLS (Splus)	ML-IRLS (SAS)
250	0.3	1.7	1.4	1.5
500	0.3	0.4	0.3	0.3
750	0.3	0.6	0.5	0.5
900	0.3	0.1	0.007	0.007
1000	0.3	0.3	0.2	0.2
5000	0.3	-0.1	-1.6	-0.2
250	2.0	2.7	1.9	1.9
500	2.0	2.0	1.6	1.6
750	2.0	1.2	0.9	0.9
1000	2.0	1.0	0.8	0.8
500	-0.3	0.1	-0.05	-0.05
500	5.0	4.5	3.3	3.3

Simulated univariate probit model with a single predictor and single binary outcome. Simulation results fit using weighted least squares with robust standard errors (WLSMV) in Mplus and maximum likelihood via iteratively re-weighted least squares (ML-IRLS) in Splus, and SAS.

Discussion

The purpose of this study was to evaluate the performance of NLMMs relative to SEMs in the analysis of mediated longitudinal binary outcomes in a setting where the SEM is presumed to be optimal. We found model performance differed based on the link function that was used in the non-linear portion of the models. Based on simulations performed across a variety of settings, the logistic NLMM performed well with respect to bias, coverage probability and power relative to the logistic SEM. The results were similar for the SEM and scaled NLMM in the logistic model setting, with both accurately estimating the effect of the time-dependent

predictor on the longitudinal binary outcome. Application of these techniques to a real-date example from a prospective cohort assessing the effect of heavy alcohol consumption on low CD4 cell count also illustrate the similarity of results from the logit SEM and NLMM.

For the probit model, however, the SEM consistently overestimated the total effects of the predictor and generally had larger bias and lower power compared to the NLMM in both mediated and non-mediated data. The larger bias may be due to the weighted least squares estimation method used for the probit SEMs (fit with Mplus), which differs from the maximum likelihood estimation method used for the NLMM (fit with SAS) and for the logistic SEM (fit with Mplus). In contrast to the SEM, the scaled probit NLMM had good performance (low bias and high power and coverage probability) with adequate sample sizes. Similar results were observed in the real data example where estimates from the probit SEM appeared larger than those from the probit NLMM.

The results showing similar estimated effects for the SEM and NLMM in the logistic model setting are similar to results seen in the non-mediated case where SEM was compared to the generalized estimating equations (a non-SEM) technique. Palta and Lin [17] compared probit models for SEM and generalized estimating equations in the analysis of data from a cohort study and found that when appropriately scaled, the two models yielded similar results. They noted, however, that the SEM allowed for more flexible specification of variance structure and therefore allowed coefficients to be scaled to provide marginal or cluster-specific interpretation.

To obtain scaled NLMM estimates in practice, it may be preferable to model the mediation by fitting separate equations, one for each pathway, using maximum likelihood rather than weighted least squares (the only estimation method currently available for probit link models in MPlus). The potential burden of fitting multiple equations separately rather than simultaneously using SEMs may be outweighed by the benefit of using maximum likelihood estimation which, in the probit model simulations, appeared to produce less biased results. In addition, estimating scale parameters and using the product of coefficients method appears to produce acceptable estimates of the total effect of the exposure. Our study demonstrated that results using this approach were similar to those obtained when NLMM results were scaled using true parameter values. If indirect effects are of interest and the NLMM is used to analyze the mediated longitudinal binary data, scaling will also be necessary. Unlike the case with linear models for continuous outcome data, the product of coefficients method is not equivalent to the "difference of coefficients" method of determining the indirect or mediated effect [24] in the case of binary outcomes. Using the difference of coefficients approach in linear models, the total effect is obtained by fitting a model that excludes the mediating variable and the direct effect is obtained by fitting a model including the mediating variable. The indirect effect is then determined by taking the difference between the total effect and direct effect. However, in the binary case, the scale of the direct effect obtained from a model that includes the mediating variable is different from the scale of the total effect obtained from a model that excludes the mediating variable [16]. As demonstrated by MacKinnon et al., to obtain comparable estimates of the indirect effect in binary outcome models, the total effect must be appropriately scaled before the difference is taken.

This study presents results based on simulated data from a single-mediator model. Conclusions from these results may not be generalizable to scenarios with different data characteristics. For example, in scenarios with multiple mediators and pathways, the advantages and disadvantages of NLMMs relative to SEMs may differ. The performance of NLMMs and SEMs in other scenarios, such as the analysis of nominal and ordinal outcomes as well as the case of multiple mediators, should be evaluated in future studies.

Conclusions

Overall, we found the NLMM performed sufficiently well in the analysis of mediated longitudinal binary outcomes with respect to bias, coverage probability, and power. Under the logistic model, both the NLMM and SEM had acceptable performance and the results for the two types of models were similar. The NLMM requires scaling of the regression parameters and this scaling requires fitting additional models to separately estimate direct effects of the predictor, and effects of the primary predictor on the mediator. An advantage of the SEM is that it can fit all of the linear and non-linear models simultaneously, avoiding the burden of fitting multiple models. For the probit model, however, the SEM estimated using weighted least squares may overestimate effects. In contrast, the NLMM appears to perform adequately across a range of settings and therefore is preferred over the SEM for probit models.

Appendix

Estimating the Scaling Factor for Total Effect Estimates

Scaling effect estimates is necessary to obtain total effect estimates that represent the total effect of the predictor on the outcome, conditional on the mediation. This is the case when one wishes to compare estimates from SEMs to estimates from NLMMs, but also necessary if one wishes to compute indirect (i.e. mediated) effects using results from NLMMs only. The difference in scale

for models with differing sets of predictors is due to the variance of the residual error term being fixed. This is not an issue in linear models where this variance is estimated rather than fixed [16-18]. In practice, the unknown parameters necessary for scaling, i.e. θ (the variance of ϵ_j and λ (the effect of the mediator on the outcome), can be estimated by fitting the following additional models:

$$M_j = \alpha_0 + \alpha_1 z_j + \varepsilon_j \quad (12)$$

$$\Phi^{-1}(Y_j) = \beta_0 + \beta_1 w + \beta_2 t_j + \beta_3 z_j + \beta_4 M_j + b_1 + b_2 t_j \quad (13)$$

The first model (Equation 12) is similar to part of the structural model in the SEM (Equation 5), but instead is fit as a general linear model for longitudinal data with M_j as the outcome and z_j as the predictor that allows correlation between the repeated observations. Using data from the alcohol and HIV example described earlier, this would be a model with longitudinal ART adherence measures as the outcome and measures of heavy alcohol consumption as time-varying predictors. The second model is an NLMM modeling Y_j as a function of a random intercept and slope, and fixed effects for the continuous covariate (w), time (t_j), the repeated binary predictor (z_j), and the repeated mediator (M_j). For the HIV example, this would be a model with low CD4 count as the outcome and include both heavy alcohol consumption and ART adherence as predictors.

Using Equations 12 and 13, the estimated variance of ϵ_j provides an estimate of θ and the estimated coefficient β_4 associated with M_j provides an estimate of λ. The original, unscaled NLMM estimate of the total effect of the primary predictor, β_3 from Equation 8), can then be rescaled by multiplying by the factor: $\sqrt{1 + \hat{\lambda}^2\hat{\theta}}$ for a probit model (or $\left[\left(\frac{\pi^2}{3} + \hat{\lambda}^2\hat{\theta}\right)/\frac{\pi^2}{3}\right]^{\frac{1}{2}}$ for a logit model).

An alternative approach to using scaled regression coefficients would be to model indirect and direct pathways separately and obtain the total effect by summing the indirect and direct effects. That is, estimates of the coefficients β_3 and β_4 associated with z_j and M_j, respectively (from Equation 13) can be used along with estimates of the coefficient α_1 associated with z_j (from Equation 12) to obtain the estimated total effect of the main independent variable $\hat{\beta}_3 + \hat{\alpha}_1\hat{\beta}_4$.

In the current simulation study, we calculated the total effects of the main independent variable using both of the approaches described above. For the probit model, both methods yielded results comparable to those where the true parameters were known. The rescaled NLMM resulted in a parameter estimate (standard error) of 0.407 (0.105) which is a bias of 1.7%. The product of coefficients method yielded an estimate of 0.400 (0.102), which is a bias of 1.7%. Both estimates were very similar to those obtained using true values for the scaling factor, parameter estimate of 0.407 (0.096) and bias of 1.7%.

Acknowledgements
The authors wish to thank Howard Cabral and Timothy Heeren for thoughtful discussions of the problem. We are also grateful to Jeffrey H. Samet, PI of the HIV-LIVE Study (NIAAA: R01-AA13216), for use of the data.

Author details
[1]Department of Biostatistics Boston University School of Public Health 801 Massachusetts Avenue 3rd Floor Boston, MA 02118 USA. [2]Clinical Research Program, Children's Hospital Boston and Harvard Medical School, 300 Longwood Avenue Boston, MA 02115 USA.

Authors' contributions
EAB was involved in the conception of the study, designing and performing the simulation analysis and drafting the manuscript. DMC was involved in conception of the study, designing the simulation analysis and drafting the final manuscript. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 15 July 2011 Accepted: 24 January 2012
Published: 24 January 2012

References

1. Muthen BO: **A General Structural Equation Model With Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators.** *Psychometrika* 1984, **49**:115-132.
2. Davidian M, Giltinan D: *Nonlinear Models for Repeated Measurement Data* New York, New York: Chapman and Hall; 1995.
3. Diggle PJ, Liang KY, Zeger SL: *Analysis of Longitudinal Data* New York, New York: Oxford University Press; 1994.
4. McArdle JJ, Epstein D: **Latent growth curves within developmental structural equation models.** *Child Development* 1987, **58**:110-133.
5. Bollen KA: *Structural Equations With Latent Variables* New York, New York: Wiley; 1989.
6. McArdle JJ, Aber MS: *Statistical Methods in Longitudinal Research* New York: Academic Press; 1990, 151-224, chap. Patterns of change within latent variable structural equations models.
7. Muthen B: *Best Methods for the Analysis of Change: Recent Advances, Unanswered Questions, Future Directions* Washington, D.C.: APA; 1991, 1-17, chap. Analysis of Longitudinal Data Using Latent Variable Models with Varying Parameters.
8. Muthen BO, Curran PJ: **General Longitudinal Modeling of Individual Differences in Experimental Designs: A Latent Variable Framework for Analysis and Power Estimation.** *Psychological Methods* 1997, **2(4)**:371-402.
9. Rovine MJ, Molenaar PCM: **A Structural Modeling Approach to a Multilevel Random Coefficients Model.** *Multivariate Behavioral Research* 2000, **35**:51-88.
10. Rovine MJ, Molenaar PCM: *New methods for the analysis of change* Washington, D.C.: APA; 2001, 67-96, chap. A Structural Equations Modeling Approach to the General Linear Mixed Model.
11. Stoel RD, van Den Wittenboer G, Hox J: **Estimating Multilevel Linear Models as Structural Equation Models.** *Metodologia de las Ciencias del Comportamiento* 2003, **5**:21-42.
12. Curran PJ: **Have Multilevel Models Been Structural Equation Models All Along?** *Multivariate Behavioral Research* 2003, **38(4)**:529-569.
13. Bauer DJ: **Estimating Multilevel Linear Models as Structural Equation Models.** *Journal of Educational and Behavioral Statistics* 2003, **28(2)**:135-167.

14. Blood EA, Cabral H, Heeren T, Cheng DM: **Performance of mixed effects models in the analysis of mediated longitudinal data.** *BMC Medical Research Methodology* 2010, **16(10)**.
15. Blood EA, Cheng DM: **The Use of Mixed Models For the Analysis of Mediated Data with Time-Dependent Predictors.** *Journal of Environmental and Public Health* 2011, 2011.
16. MacKinnon D, Lockwood C, Brown C, Wang W, Hoffman J: **The intermediate endpoint effect in logistic and probit regression.** *Clinical Trials* 2007, **4**:499-513.
17. Palta M, Lin CY: **Latent Variables, Measurement Error and Methods for Analysing Longitudinal Binary and Ordinal Data.** *Statistics in Medicine* 1999, **18**:385-396.
18. Fitzmaurice GM, Laird NM, Ware JH: *Applied Longitudinal Analysis* Hoboken, New Jersey: John Wiley and Sons, Inc; 2004.
19. Muthen L, Muthen B: *Mplus User's Guide.* Fourth edition. Los Angeles, CA: Muthen & Muthen; 1998, chap. 3.
20. Rosner B: *Fundamentals of Biostatistics* Pacific Grove, California: Dubxbury Press; 2005.
21. SAS Institute Inc: *SAS 9.2 Help and Documentation* Cary, NC: SAS Institute Inc; 2002.
22. Cohen J: *Statistical Power Analysis for the Behavioral Sciences* Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
23. Samet J, Cheng D, Libman H, Nunes D, Alperen J, Saitz R: **Alcohol consumption and HIV disease progression.** *Journal of Acquired Immune Deficiency Syndrome* 2007, **46(2)**:194-199.
24. MacKinnon DP, Warsi G, Dwyer JH: **A Simulation Study of Mediated Effect Measures.** *Multivariate Behavioral Research* 1995, **30**:41-62.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2288/12/5/prepub

doi:10.1186/1471-2288-12-5
Cite this article as: Blood and Cheng: **Non-linear mixed models in the analysis of mediated longitudinal data with binary outcomes.** *BMC Medical Research Methodology* 2012 **12**:5.

The Role of Patients' Explanatory Models and Daily-Lived Experience in Hypertension Self-Management

Barbara G. Bokhour, PhD[1,2], Ellen S. Cohn, ScD[5], Dharma E. Cortés, PhD[6,12], Jeffrey L. Solomon, PhD[1], Gemmae M. Fix, PhD[1], A. Rani Elwy, PhD[1,2], Nora Mueller, MA[1], Lois A. Katz, MD[7,8], Paul Haidet, MD, MPH[9], Alexander R. Green, MD, MPH[10,11,12], Ann M. Borzecki, MD, MPH[1,4], and Nancy R. Kressin, PhD[3,4]

[1]Center for Health Quality, Outcomes & Economic Research, ENRM Veteran Affairs Medical Center, Bedford, MA, USA; [2]Boston University School of Public Health, Boston, MA, USA; [3]VA Boston Healthcare System, Boston, MA, USA; [4]Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA; [5]Department of Occupational Therapy, Boston University, College of Health and Rehabilitation Sciences, Boston, MA, USA; [6]Cambridge Health Alliance, Cambridge, MA, USA; [7]VA New York Harbor Healthcare System, New York, NY, USA; [8]New York University School of Medicine, New York, NY, USA; [9]Office of Medical Education and the Departments of Medicine, Humanities, and Public Health Sciences, The Pennsylvania State University College of Medicine, Hershey, PA, USA; [10]Massachusetts General Hospital, Mongan Institute for Health Policy, Boston, MA, USA; [11]Massachusetts General Hospital, Disparities Solutions Center, Boston, MA, USA; [12]Harvard University School of Medicine, Boston, MA, USA.

BACKGROUND: Uncontrolled hypertension remains a significant problem for many patients. Few interventions to improve patients' hypertension self-management have had lasting effects. Previous work has focused largely on patients' beliefs as predictors of behavior, but little is understood about beliefs as they are embedded in patients' social contexts.

OBJECTIVE: This study aims to explore how patients' "explanatory models" of hypertension (understandings of the causes, mechanisms or pathophysiology, course of illness, symptoms and effects of treatment) and social context relate to their reported daily hypertension self-management behaviors.

DESIGN: Semi-structured qualitative interviews with a diverse group of patients at two large urban Veterans Administration Medical centers.

PARTICIPANTS (OR PATIENTS OR SUBJECTS): African-American, white and Latino Veterans Affairs (VA) primary care patients with uncontrolled blood pressure.

APPROACH: We conducted thematic analysis using tools of grounded theory to identify key themes surrounding patients' explanatory models, social context and hypertension management behaviors.

RESULTS: Patients' perceptions of the cause and course of hypertension, experiences of hypertension symptoms, and beliefs about the effectiveness of treatment were related to different hypertension self-management behaviors. Moreover, patients' daily-lived experiences, such as an isolated lifestyle, serious competing health problems, a lack of habits and routines, barriers to exercise and prioritizing lifestyle choices, also interfered with optimal hypertension self-management.

CONCLUSIONS: Designing interventions to improve patients' hypertension self-management requires consideration of patients' explanatory models and their daily-lived experience. We propose a new conceptual model — the dynamic model of hypertension self-management behavior — which incorporates these key elements of patients' experiences.

Received October 11, 2011
Revised April 11, 2012
Accepted June 5, 2012
Published online July 21, 2012

KEY WORDS: hypertension; medication adherence; qualitative methods; health behavior; self-management.
J Gen Intern Med 27(12):1626–34
DOI: 10.1007/s11606-012-2141-2

INTRODUCTION

Hypertension continues to be a major US health problem; it is underdiagnosed, undertreated, and even when treated, often poorly controlled; over one-third of patients on antihypertensive medications have uncontrolled disease.[1] Poor hypertension control has been attributed to provider factors, such as clinical inertia in prescribing adequate medications[2,3] and patient factors, including poor adherence to medications and lifestyle management.[4–6] Few interventions to improve hypertension control have had resounding, sustained effects.[7]. Thus, finding ways to improve hypertension self-management remains a significant challenge.

Prior interventions have included educating patients about hypertension, improving BP self-monitoring at home, and nurse or pharmacist-led care. A recent Cochrane review found that purely educational interventions improved patient knowledge but were largely ineffective in improving hypertension control. Patient self-monitoring and nurse or pharmacist-led interventions held greater promise, although results were heterogeneous.[7]

One reason for the failure of purely educational interventions may be that they lack a patient-centered focus, thereby failing to address patients' unique barriers to hypertension

self-management. A patient-centered approach would incorporate how patients understand hypertension, attempt to manage hypertension, and prioritize its management in their lives. Many interventions are built upon theories that beliefs or illness perceptions drive patients' illness self-management.[8,9] Kleinman[10] reframes beliefs as integrated "explanatory models (EMs)" — patients' understanding of the causes, pathophysiology, course of illness, symptoms, and effects of treatment. EMs of an illness like hypertension are situated within social contexts; they are formed and persist in the context of the communities in which patients live. Thus, in order to influence health-related self-management behaviors, one must understand both patients' beliefs about an illness and their social context.

Further understanding hypertension self-management experiences would foster a patient-centered approach to hypertension control. Previous studies have examined patients' 'lay' models of hypertension,[11–13] but not how EMs relate to social context, or how they impact various hypertension self-management behaviors. In this qualitative study, we extend prior work by examining patients' EMs, social contexts, and hypertension self-management behaviors. We explore these aspects of patients' hypertension experiences and propose a new conceptual model of patients' illness self-management.

METHODS

Participants and Recruitment

We recruited a purposive sample of African-American, white and Latino patients with uncontrolled hypertension from primary care clinics at two large US Veterans Affairs (VA) Medical Centers serving diverse populations located in the Northeast and Southwest, respectively. In a 1-year period, we sought to recruit 20 patients from each racial/ethnic group to obtain a wide range of responses, with the expectation that we would reach thematic saturation for a diverse population. Each site's institutional review board approved the study.

Patients were eligible to participate if they had a diagnosis of hypertension and uncontrolled blood pressure (BP) (≥140/90 mmHg) documented at least once in the previous 6 months in the VA's electronic medical record, and a BP>140/90 mmHg at their primary care visit. A research assistant (RA) called patients prior to their appointment to invite participation. At the visit, the RA administered eligibility screening, collected demographic information, took BP readings and obtained written informed consent. The data were collected between February 2007 and June 2008.

Data Collection

A trained RA conducted semi-structured 1.5 hour interviews with patients. A $20.00 gift card was given for participation. Patients were asked to describe their family, a typical day, experiences living with hypertension, knowledge about hypertension, what they thought caused hypertension, the severity of their hypertension, their concerns about hypertension, communication with their provider, and medications and strategies they used to manage their hypertension.

Data Analysis

Interviews were audio-recorded, professionally transcribed verbatim, and reviewed for accuracy. We used qualitative analytic techniques informed by grounded theory.[14,15] We started by exploring how EMs affected hypertension self-management. Consistent with a grounded theory approach, we sought to identify other aspects of patient experience emergent in the data. Five team members began by open coding five transcripts together, identifying themes grounded in participants' language, and generating a coding dictionary that was refined with coding of subsequent interviews. We collapsed codes into broader categories, informed by the data and EM categories of "cause," "course," "pathophysiology," "symptoms," and "treatment," and by aspects of social context. We identified participants' hypertension self-management behaviors — reports of what they did to control their hypertension — and developed summary templates describing each participant's EMs, social context, hypertension self-management behaviors, perceptions of clinical encounters, and primary barrier to hypertension self-management. Investigators paired up to analyze subsequent interviews and complete templates, which were reviewed by the entire team, and cases were contrasted using constant comparison analysis. Throughout the analytic process, we iteratively refined conceptual links between constructs to develop a conceptual model.

RESULTS

We interviewed 48 patients (19 African-American, 20 white, and 9 Hispanic). We failed to recruit the target number of Hispanic patients, due to the relatively low Hispanic prevalence at our sites and greater participation refusal. The average age was 60; 92 % were men; 88 % had at least a high school degree, and yet income was low (see Table 1). The average BP at the clinic visit was 160/93 mmHg. We identified two major domains affecting patients' hypertension self-management behaviors: 1) four different aspects of EMs, including beliefs regarding the cause of hypertension, hypertension symptoms, the illness course (including whether they thought hypertension was chronic or intermittent), and treatment value and effects; and 2) aspects of patients' social context, routines, habits, and

Table 1. Participant Demographics

Interview participants (N=48)	
Males	44
Age	Mean – 60 (SD – 10.31)
Race	
Black	20
White	19
Latino	9
Education	
≤ 8th grade	3
Some high school	2
Completed high school	14
Some college	8
Completed college	8
Graduate school	2
Unknown	1
Income ($)	
<15,000	14
15,000 – 20,000	7
20,000 – 40,000	13
40,000 – 60,000	5
>80,000	2
Unknown	7

competing health problems, which we termed "daily-lived experience" (DLE).

Patients described a wide range of EMs and DLEs that adversely affected their hypertension self-management behaviors and BP control (Tables 2 and 3). We first treat EM and DLE as separate and provide exemplary quotes demonstrating how EMs and DLEs were related to patients' reported hypertension self-management behaviors. We then show how EMs and DLEs may relate to one another in shaping self-management, and describe the resulting conceptual model. For 47 of the 48 patients we interviewed, we identified at least one EM or DLE that adversely affected their hypertension self-management behavior. We did not identify any unique racial/ethnic differences in EMs or DLE.

Explanatory Models (EMs)

One aspect of EMs was notably absent from patients' interviews: pathophysiology, or a biomedical conceptualization of the condition. Few patients described their understanding of the pathophysiology of hypertension, and when they did, these understandings were not linked to hypertension self-management behaviors. In contrast, perceptions of cause, illness course, symptoms, and treatment were related to self-management in several ways.

Cause. When asked what caused their high BP, patients' responses corresponded to notions of both the chronic disease of hypertension and to a temporary increase in BP associated with daily activities. Patients did not distinguish these two 'illnesses,' and often engaged in behaviors to avoid the latter, rather than behaviors that addressed biomedically attributed causes of hypertension . Patients noted causes including: 1) heredity, 2) general life stress, 3) poor diet (especially salt intake), and 4) lack of exercise.

Table 2. Explanatory Models and Corresponding Hypertension Self-Management Behaviors. These Represent the Range of Explanatory Models Described by Patients in Qualitative Interviews as Affecting Hypertension Self-Management Behaviors

Explanatory Model	Hypertension Self-Management Behavior
Cause	
Stress as primary cause	Stays calm, avoids stressful situations; takes antidepressant as treatment
Exercise causes increased BP	Avoids exercise to keep BP low
Pain causes increased BP	Managing pain, taking pain medications will control hypertension
Course of Illness	
Hypertension comes and goes	Takes medications when BP goes up
BP cannot be controlled	Won't exercise and forgets medications
Little concern about hypertension – does not affect my life	Avoids going to the doctor; Forgets medications
Own definition of what is considered 'high'	Only take medication when BP is >190/100 mmHg
Symptoms	
I can tell when my BP is high; I get headaches, dizziness when my BP is high	Takes medications only when symptoms occur
I have no symptoms of high blood pressure; therefore it is not a problem	Doesn't take medications
Eating bacon doesn't make me feel bad, so it doesn't affect my BP.	Eats bacon as desired
Treatment	
Only exercise can help me control my HBP	Exercises and therefore allows himself to smoke, drink and not take medication.
Garlic and vinegar can help me control my HBP.	Focuses on these remedies, while not taking medications or altering diet or sodium intake

Patients also described causes typically associated with temporary rises in BP: 1) increased situational stress, 2) pain, and 3) exercising. Hypertension self-management behaviors were related to both types of perceived causes. For example, one woman who responded that stress was the cause of her hypertension went on to say:

Table 3. Daily-Lived Experience and Corresponding Reported Behaviors that Interfere With Hypertension Self-Management. These Represent the Range of DLEs Described by Patients in Qualitative Interviews as Affecting Behaviors that Might Interfere with Hypertension Self-Management

Daily-Lived Experience	Behaviors Interfering with Hypertension Self-Management
Lonely, isolated lifestyle	Eats out at restaurants for social interaction, which limits ability to control salt intake.
Serious competing chronic illnesses	Disregards hypertension in favor of managing these illnesses. Unable to engage in exercise.
Lack of routines and an unstructured lifestyle	Forgets to take medications; misses doctor appointments.
Experiences medication side effects, such as impotence	Decides not to take medications.
Frequent alcohol use	Avoids medications while drinking, doubles up on medications when not drinking.

> "I think that my main reason for moving upstate is to find tranquility and maybe calm myself. And I keep saying, if I move upstate, I won't have hypertension, I probably will not need medication, the lifestyle's going to be different, it'll be more relaxed." (#1– Latino woman, age 51)

This woman had described a complex, stressful city life, including work and family stressors. Her EM that stress was the cause of her hypertension, reflected in her statement that if she moved she would no longer have hypertension, led her to consider altering her living situation. This view could perhaps contribute to poorer medication adherence.

Course of Illness. The course of illness refers to patients' perceptions of the path, controllability and severity of the illness. Some patients viewed hypertension as an intermittent problem that comes and goes. They were not concerned with hypertension as a chronic problem that needed constant management; rather, they focused on it when they knew their BP was high. For example, one woman stated:

> "If my blood pressure is anywhere between, let's say, 190 over 100, then I'm okay. As far as I'm concerned, I feel fine….right now that's what it's running about. And I'm still taking my medicine, though I'm not up to the max dose that I was up to before. So in my eyes, I'm fine. … if after a point … it starts creeping up to 220 over 120, then I have to do something at that point. It's getting too high. I know that. But if it's in the low 90, 180 range, to me, that's not considered high for me. Because I feel I'm still feeling fine. I can function." (#2 – African American woman, age 53)

This woman expressed an EM of hypertension as intermittent; moreover, 'normal BP' for her is incongruent with the biomedical view of poor hypertension control. Subsequently, she altered her self-management behavior depending on her EM of 'high' blood pressure. She hinted that how she feels and functions, rather than an objective measurement, changes her behavior.

Symptoms. Patients discussed symptoms of hypertension in two ways: 1) "There are no symptoms of hypertension," or 2) "I can feel when my blood pressure is high" (see the previous quote). In the first view, patients stated that since they did not have any symptoms associated with their hypertension, there was no need to adhere to hypertension management recommendations. One man who didn't have symptoms stated:

> "[Hypertension] doesn't affect me. If it would just hit me, "boom!" and it would like put me to a point where I was … I'm getting dizzy from eating bacon or whatever, then I would change that, I wouldn't do that. But it hasn't." (#3 – African American man, age 73)

He expressed an EM in which the presence of symptoms indicated a need for concern, incongruous with the view of hypertension as the silent killer. Subsequently, he continued to eat bacon, despite recommendations to the contrary.

Other patients described hypertension symptoms including headaches, "pressure on your chest," feeling "queasy or uneasy," neck aches, backaches, or racing pulse. One man described having a headache and feeling "the blood [rush] to your head." Many of these patients altered their hypertension self-management behaviors depending on their symptoms. One man stated:

> "I can almost tell when it's a little bit getting higher, I start feeling a little different in my head, you know, it's like a headache or something, like a little pressure. So I back off on eating (inaudible) and whatnot, or I take an extra pill on that day. I mean, I usually take two, and I would do that and it sort of levels it again. Or I really stop eating everything that has salt and sodium … and that seems to work." (#4 – Latino man, age 63)

The EM that high blood pressure produces symptoms led him to alter his medication and diet depending upon his symptoms. This behavior, in turn, may contribute to poorly controlled BP.

Treatment. Beliefs about treatment effectiveness shaped hypertension self-management, and were linked to patients' perceptions of causes and symptoms. Some endorsed taking medications, altering diet and engaging in exercise as effective treatment. Others, such as the following man, expressed concerns about taking medications for BP control, and altered treatment accordingly.

> "Higher dosages concern me. I felt that my pressure was higher than…. I must have done something and it started getting higher, and I came to your walk-in clinic. And I was given a higher dosage. And I didn't take it, I just went home and I relaxed. I think I might have taken one actually, or one-half." (#5 – Latino man, age 63)

This man's EM that high medication doses may be harmful led him to rely more on relaxation than taking medications to control his hypertension. Note this behavior also reflects an EM that relaxation is an effective treatment.

Another man stated that he didn't like taking medication, liked to drink and smoke, and that only exercising would control his hypertension:

> "So, like I run. I run, until I, you know, I'm tired and can't run any more. I stop. I sit. I collect myself and then I go again. … I smoke, so I work on my lungs as much as I can. And the thing is I'm trying to clear them so I can breathe better and basically have more wind… But now, I'm on—they got me on—blood pressure pills. I'm taking medication for high blood pressure."
> *INTERVIEWER:"Right. And how many are you on for blood pressure?"*
> "Just—I take one a day. You know, sometimes I do. Sometimes I miss it, you know. But my doctor, my primary care doctor, she wants me to take these every day. I don't like medication too much. I like natural things." (#6 – African American man, age 64)

His EM was that taking medication was unnatural, and therefore concerning. Despite his acknowledgement that smoking is bad for him, he expressed an EM that getting exercise can overcome negative effects from smoking. He subsequently relied solely on exercise to control his BP.

Daily-Lived Experience (DLE)

Patients were asked to tell the interviewer about their lives and describe a typical day. The ways patients managed their hypertension were often embedded in descriptions of their daily activities. We identified five aspects of patients' DLE that were related to hypertension self-management behaviors: 1) isolated lifestyle, 2) serious competing health problems, 3) lack of habits and routines, 4) barriers to exercise, and 5) prioritizing lifestyle choices (Table 3). Below we provide some exemplars, noting also how patients' EMs intersected with DLE.

Lonely and Isolated Lifestyle. Many patients lived alone, were retired, unemployed, or on disability. Subsequently, they had little regular daily social interaction. One exception was venturing out to a local diner for daily meals, through which they attained the social interaction necessary to sustain their mental health. This practice, however, meant that they had little control over the amount of salt in their food. One man stated:

> "I live alone, you know? So I don't like cooking for one person. So I always go someplace to get lunches or dinners or whatever, or I go out and eat. Even eggs, you know, you can have a pinch of sodium in eggs, they already have it; they have sodium already. So everything that I eat might have sodium, and that little bit could bring it way up". (#7 – Latino man, age 61)

This man's EM was that sodium was bad for his hypertension. However, his DLE included eating at restaurants for social engagement, interfering with his ability to control his sodium intake.

Serious Competing Health Problems. Most participants had multiple co-occurring illnesses that competed with their hypertension self-management. Some confused hypertension treatment regimens with regimens for other conditions. Problems such as having arthritis or pain interfered with following exercise recommendations. For others, controlling hypertension was not considered important relative to other conditions. For example, when asked about how concerned he was about hypertension, one man who was HIV positive and had diabetes responded:

> "I'm not at all concerned about it, because I got the HIV and I got the diabetes. And high blood pressure is just another one."
> *INTERVIEWER: "How has it affected your overall health, if in any way?"*
> "I think I keep saying, to the best of my knowledge, I don't feel it anywhere." (#8 – African American man, age 67)

The DLE of having other competing health problems to manage interfered with this man's hypertension self-management. His EM of hypertension as symptomless further rendered it unimportant relative to his other health problems.

Lack of habits and routines. Some patients described highly unstructured lifestyles with few routines, and forgetting to take medications. When asked by the interviewer what made it difficult for him to lower his BP, one man replied:

> "Some very simple things. First of all, sometimes I scrap my morning routine. Sometimes I get up in the morning. By the time I realize, 'Oh, I didn't take my medicine yet,' it's one or two in the afternoon, okay." (#9 – white man, age 54)

This man's DLE included disrupted routines, which subsequently resulted in not taking his medication as prescribed. This disruption, rather than any EM about his medications, interfered with medication adherence and hypertension control.

Others had unpredictable daily lives. One man described working on a boat at sea, during which time he often forgot

to bring enough medication, was unable to obtain refills, and sometimes forgot to take it. He also described difficulty controlling his diet, as the ship's cook prepared very salty food. Thus, despite having medication and limiting salt as part of his EM for controlling BP, his daily life presented barriers to doing so.

Barriers to Exercise. Many participants expressed an EM that exercise was important to managing their hypertension, but they experienced barriers to exercising. One barrier was other health conditions, such as arthritis, that prevented them from engaging in exercise. Another was an unsafe environment. One man explained:

> "I do the exercise. Sometimes I go out and I walk up and down the steps. …'You do that for 30 minutes,' I said, 'It is almost equivalent to walking a mile' … because in my neighborhood, it is not real safe to be walking around, up and down the street."
> *INTERVIEWER: "They got loose dogs everywhere, huh?"*
> "Not only the dogs, gangsters." (#10 – white man, age 64)

This man's DLE of living in an unsafe neighborhood challenged his ability to get his desired exercise, despite the fact that he endorsed an EM in which exercise helps control BP.

Prioritizing Lifestyle Choices. Patients described lifestyle choices that affected their hypertension self-management. For example, one man stated that he "liked to drink."

> "Sometimes I wouldn't take [BP medication] for two or three days because I'm drinking. And when I continue, I'll take two for a couple days, and that's it. You know, if I feel my pressure is too high so I just double-dose, you get it? Only for a couple days, and then I go back to the regular dose." (#11 – African American man, age 57)

Drinking alcohol was part of his DLE. He understood the potential problems of drinking while taking his medication, so he altered his medication use. He also expressed an EM of the effects of treatment: the total amount of medication taken in a week is important, rather than daily adherence.

EMs and DLE Together

Patients' narratives about their hypertension reveal the intertwined nature of EMs and DLEs. In some instances, the EM might drive DLE and behavior. For example, in quote #5, the gentleman's EM is that he doesn't wish to take medications because they aren't natural. This subsequently drives his behavior around exercising, which is something that he *can* do in his daily life.

In other cases, DLE might drive the EM. In quote #1, the woman described a DLE in which stress was prevalent in the inner city context in which she lived. She attributed her high BP to this stressful life context, and in turn, believed that moving to a more peaceful setting would control her BP. If she moved, and her BP did not improve, this could alter her EM about the cause of her hypertension.

Conversely, there are instances where patients' DLE interfered with their abilities to act on their EM. This is especially salient in the case of sodium intake. Many patients reported that lowering their salt intake was important for controlling their BP, but as noted in quote #7, eating out and eating prepared foods limited patients' abilities to control their salt intake. This in turn may reinforce an EM that excess salt causes hypertension.

Conceptual links between these constructs resulted in the development of the Dynamic Model of Chronic Disease Self-Management (see Fig. 1). In this model, EMs and DLE continually shape one another in relationship to behavior, and may change over the course of a patient's illness, as has been noted in our childhood asthma work.[16] When discussing EMs and DLEs, patients suggested that communication with providers was also important to their self-management. Research supports the importance of this dimension on patient adherence.[17] Thus, we also include patient-provider communication in the model.

DISCUSSION

Patients' reports of their EMs and DLEs demonstrate that hypertension self-management does not occur primarily in the doctor's office. Rather, patients self-manage their hypertension in complex social contexts, in which they develop hypertension EMs while engaging in daily activities. Our findings show different ways in which EMs and DLEs can affect patients' hypertension self-management, including adherence to medications and recommended healthy behaviors.

Previous studies have demonstrated how differences in EMs affect patients' illness management behaviors[13,18] and confirm our finding that patients' EMs lead patients toward particular hypertension self-management behaviors that may interfere with hypertension control.[19] 12. Huertin-Roberts[12] found that African American women differentiated between "high blood pressure," thought to be responsive to diet and medication changes, and "high-pertension," a disease of the nerves caused by stress and factors that elicit emotional upset, and responsive to stress

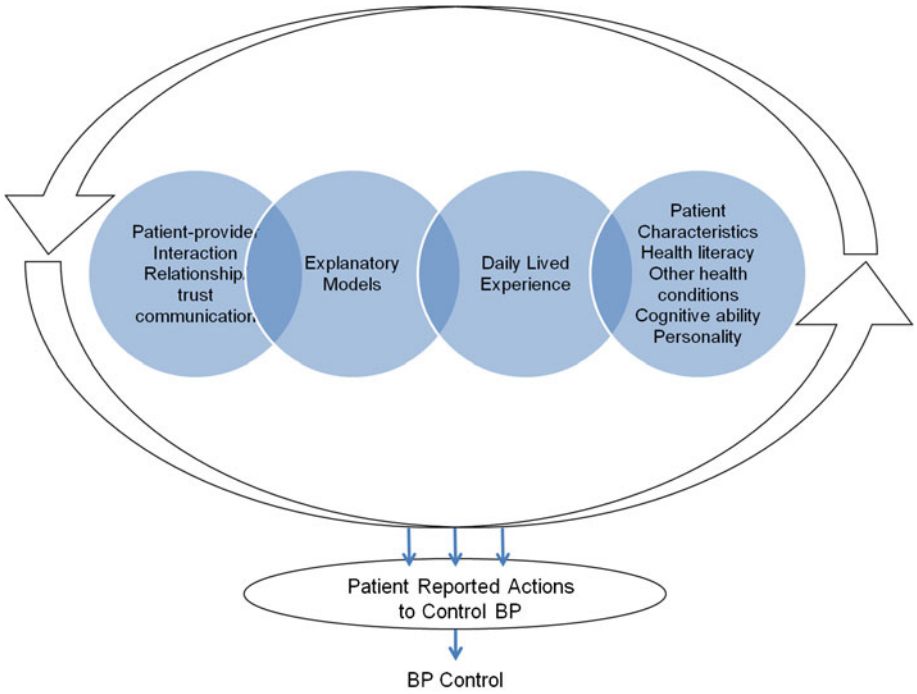


Figure 1. Dynamic model of hypertension self-management behavior.

relief rather than antihypertensives. Others have found that medication beliefs contribute to medication adherence.[9,20] These studies are consistent with theories of health behavior that account for behavior in terms of patients' internal cognitive assessments and motivations for engaging in healthy behaviors.[21–24]

Our findings demonstrate, however, that patients' self-management is based on more than their EMs or beliefs. Patients' DLE, including their social context, routines and management of competing health demands, may conflict with their EMs, and interfere with good self-management. Previous work in understanding chronic disease self-management has similarly found an important role of routines in disease self-management.[16,25]

The intervention challenge is that patients with uncontrolled hypertension have complex constellations of EMs, DLEs, other illnesses and competing priorities that converge to influence self-management behaviors. As shown in the conceptual model, DLEs and EMs are inextricably intertwined to shape behavior. If one were to imagine that an individuals' hypertension self-management experience were a strand of yarn, each fiber may constitute an EM or DLE; but if one were to pull on one strand, all the others would follow, altering the yarn altogether.

Limitations. Our findings are limited by three factors. Because this study was conducted in the VA, findings may not generalize to non-veteran populations or women, since most participants were men. Additionally, we did not observe the behaviors described; patients' reported behaviors may differ from actual behavior. Finally, although we reached saturation with our white and African American samples, we were unable to enroll as many Latino patients as intended; therefore, we may have missed some additional novel EMs or DLEs.

Interestingly, participants' infrequently stated understanding of hypertension pathophysiology was not related to self-management behaviors, suggesting that focusing on explaining how hypertension works in the body may not be an effective approach to improving hypertension control. Hypertension discussion in clinical encounters has often been restricted to adherence assessment (although even this may not occur[26]), with limited time spent counseling patients about hypertension self-management and lifestyle.[27]

While prior interventions have been tailored to particular patients' needs,[7,8,28] these interventions remain focused on providing information or education to patients to encourage them to improve their hypertension self-management behaviors, with little focus on patients' EMs in the daily life context in which they manage their illness. Many participants were able to state what they *should* do to control hypertension, but persisted in not doing so. Thus, these results suggest that improving patients' hypertension self-management requires an approach beyond assessment and education.

Encouraging patients to discuss their EMs and DLEs can enhance clinician understanding of patients' lived experience

of hypertension and subsequent self-management behaviors.[29] Some participants expressed EMs that are inconsistent with biomedical knowledge leading to poor self-management. In these instances, providers should acknowledge these differences and tailor their intervention to accommodate those beliefs.[29,30] For example, a provider may acknowledge that while reducing stress may be helpful for health, hypertension control requires additional lifestyle and medication management. Simply reminding patients to take medications and engage in healthy behavior is unlikely to be effective without addressing aspects of the patients' context. Techniques such as motivational interviewing may be effective in focusing on these aspects of hypertension management.[31] Cooper's suggestion[28] that providers need to 'get to know the patient as a person,' would likely lead providers to a better understanding of patients' DLE. These steps might lead to a decrease in contextual errors in treating hypertension: treatment errors based on a misunderstanding of the patients' context of disease management.[32]

Our findings underscore the importance of finding novel approaches in health care delivery systems to address the complex social and behavioral aspects of chronic disease self-management. While providers may address patients' EMs and DLEs during routine clinical encounters, short clinic visits often preclude such in-depth conversations, although Haidet convincingly argues that such conversations need not necessarily take more time.[33] Approaches using team based care,[28] including health coaches[34] in the context of a patient-centered medical home, may be ideal for addressing these key aspects of patient hypertension self-management.

Acknowledgements: *This study was supported by the Department of Veterans Affairs, Health Services Research and Development grant #IIR 05-062. Dr. Kressin is supported by a Senior Research Career Scientist award from the Department of Veterans Affairs, Health Services Research & Development Service (RCS 02-066-1). The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs. A version of this paper was presented as a poster at the Society of General Internal Medicine in May 2009 and the VA Health Services Research and Development annual meeting in February 2009.*

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*

Corresponding Author: *Barbara G. Bokhour; PhD; Center for Health Quality, Outcomes Economic Research, ENRM Veteran Affairs Medical Center, Bedford, MA 01730, USA (e-mail: Barbara.Bokhour@va.gov).*

REFERENCES

1. **Ostchega Y, Yoon SS, Hughes J, Louis T.** Hypertension awareness, treatment, and control–continued disparities in adults: United States, 2005–2006. NCHS Data Brief. 2008;3:1–8.
2. **Berlowitz DR, Ash AS, Hickey EC, , et al.** Inadequate management of blood pressure in a hypertensive population. N Engl J Med. 1998;339 (27):1957–1963.
3. **Rose AJ, Berlowitz DR, Orner MB, Kressin NR.** Understanding uncontrolled hypertension: is it the patient or the provider? J Clin Hypertens (Greenwich). 2007;9(12):937–943.
4. **Bosworth HB, Oddone EZ.** A model of psychosocial and cultural antecedents of blood pressure control. J Natl Med Assoc. 2002;94(4):236–248.
5. **Di Matteo M.** Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Medical Care. 2004;42(3):200–209.
6. **Stockwell DH, Madhavan S, Cohen H, Gibson G, Alderman MH.** The determinants of hypertension awareness, treatment, and control in an insured population. Am J Public Health. 1994;84(11):1768–1774.
7. **Glynn L, Murphy A, Smith S, Schroeder K, Fahey T.** Interventions used to improve control of blood pressure in patients with hypertension. Cochrane Database of Systematic Reviews. 2010;2010(3):CD005182.
8. **Bosworth HB, Olsen MK, Neary A, et al.** Take Control of Your Blood Pressure (TCYB) study: a multifactorial tailored behavioral and educational intervention for achieving blood pressure control. Patient Educ Couns. 2008;70(3):338–347.
9. **Kressin NR, Wang F, Long J, et al.** Hypertensive patients' race, health beliefs, process of care, and medication adherence. J Gen Intern Med. 2007;22(6):768–774.
10. **Kleinman A, Eisenberg L, Good B.** Culture, illness and care: Clinical lessons from anthropologic and cross-cultural research. Ann Intern Med. 1978;88(2):251–258.
11. **Blumhagen D.** The meaning of Hyper-tension. In: Chrisman NaM, TW, ed. Clinically applied anthropology: D. Reidel Publishing Company; 1982:297-323.
12. **Heurtin-Roberts S, Reisen E.** The relation of culturally influenced lay models of hypertension to compliance with treatment. American Journal of Hypertension. 1992;5(11):787–792.
13. **Schlomann P, Schmitke J.** Lay beliefs about hypertension: an interpretive synthesis of the qualitative research. J Am Acad Nurse Pract. 2007;19(7):358–367.
14. **Strauss AL.** *Qualitative analysis for social scientists*. Cambridge [Cambridgeshire]; New York: Cambridge University Press; 1987.
15. **Charmaz K.** *Constructing grounded theory*. Thousand Oaks: Sage; 2006.
16. **Yinusa-Nyahkoon LS, Cohn ES, Cortes DE, Bokhour BG.** Ecological barriers and social forces in childhood asthma management: examining routines of African American families living in the inner city. J Asthma. 2010;47(7):701–710.
17. **Rao JK, Anderson LA, Inui TS, Frankel RM.** Communication interventions make a difference in conversations between physicians and patients - A systematic review of the evidence. Medical Care. 2007;45 (4):340–349.
18. **Bokhour BG, Cohn ES, Cortes DE, et al.** Patterns of concordance and non-concordance with clinician recommendations and parents' explanatory models in children with asthma. Patient Educ Couns. 2008;70 (3):376–385.
19. **Schlomann P, Schmitke J.** Lay beliefs about hypertension: An interpretive synthesis of the qualitative research. J. Am. Acad. Nurse Pract. 2007;2007(19):358–367.
20. **Horne R, Clatworthy J, Polmear A, Weinman J.** Do hypertensive patients' beliefs about their illness and treatment influence medication adherence and quality of life? J. Hum. Hypertens. 2001;15(Suppl 1): S65–S68.
21. **Harrison JA, Mullen PD, Green LW.** A Meta analysis of studies of the health belief model with adults. Heal Educ Res. 1992;7(1):107–116.
22. **Figueiras M, Marcelino DS, Claudino A, Cortes MA, Maroco J, Weinman J.** Patients' illness schemata of hypertension: The role of beliefs for the choice of treatment. Psychol Heal. 2010;25(4):507–517.
23. **Leventhal H, Leventhal EA, Contrada RJ.** Self-regulation, health, and behavior: A perceptual-cognitive approach. Psychol Heal. 1998;13 (4):717–733.
24. **Rosenstock IM.** Why people use health services. Milbank Memorial Fund Quarterly-Health and Society. 1966;44(3):94–127.
25. **Fiese BH, Wamboldt FS, Anbar RD.** Family asthma management routines: connections to medical adherence and quality of life. J Pediatr. 2005;146(2):171–176.
26. **Bokhour BG, Berlowitz DR, Long JA, Kressin NR.** How do providers assess antihypertensive medication adherence in medical encounters? J Gen Intern Med. 2006;21(6):577–583.

27. **Bell RA, Kravitz RL.** Physician counseling for hypertension: what do doctors really do? Patient Educ Couns. 2008;72(1):115–121.
28. **Cooper LA.** A 41-year-old African American man with poorly controlled hypertension: review of patient and physician factors related to hypertension treatment adherence. JAMA. 2009;301 (12):1260–1272.
29. **Street RL Jr, Haidet P.** How well do doctors know their patients? Factors affecting physician understanding of patients' health beliefs. J Gen Intern Med. 2011;26(1):21–27.
30. **Frosch DL, Elwyn G.** I believe, therefore I do. J Gen Intern Med. 2011;26 (1):2–4.
31. **Knight KM, McGowan L, Dickens C, Bundy C.** A systematic review of motivational interviewing in physical health care settings. Br J Heal Psychol. 2006;11:319–332.
32. **Weiner SJ, Schwartz A, Weaver F, , et al.** Contextual errors and failures in individualizing patient care: a multicenter study. Ann Intern Med. 2010;153(2):69–75.
33. **Haidet P, Paterniti DA.** "Building" a history rather than "taking" one: a perspective on information sharing during the medical interview. Arch Intern Med. 2003;163(10):1134–1140.
34. **Bodenheimer T, Laing BY.** The teamlet model of primary care. Ann. Fam. Med. 2007;5(5):457–461.

Performance Improvement

Is Development of Postoperative Venous Thromboembolism Related to Thromboprophylaxis Use? A Case-Control Study in the Veterans Health Administration

Ann M. Borzecki, MD, MPH; Andrew J. Cowan, MD; Marisa Cevasco, MD, MPH; Marlena H. Shin, JD, MPH; Michael Shwartz, PhD; Kamal Itani, MD; Amy K. Rosen, PhD

Postoperative pulmonary embolism/deep vein thrombosis (pPE/DVT) is a common, costly, and potentially life-threatening complication.[1] Without thromboprophylaxis, DVT occurs in 15% to 60% and PE in 0.5% to 5% of major surgery patients, with the highest risk following orthopedic procedures, particularly total hip replacement[1]; pPE/DVT is associated with an excess of $21,000 in hospital charges, 5.4 hospital days, and fourfold 30-day mortality.[2,3]

Multiple randomized trials and meta-analyses have shown that prophylaxis, particularly pharmacoprophylaxis, can significantly reduce this risk, with reductions of almost 70% among general surgical patients.[1,4,5] Despite this, and the long-standing existence of national evidence-based PE/DVT prevention guidelines,[1,6] observational studies continue to report prophylaxis underuse.[7–9]

In recognition of these gaps in care, The Joint Commission and the Surgical Care Improvement Project (SCIP) have developed chart-based process measures addressing appropriate thromboprophylaxis use.[10] The Agency for Healthcare Research and Quality (AHRQ) has also included pPE/DVT as a Patient Safety Indicator (PSI). PSIs are outcome measures that use administrative data to identify potentially preventable adverse events.[11] The Centers for Medicare & Medicaid Services (CMS) has subsequently added the SCIP Venous Thromboembolism (VTE) measures and the pPE/DVT PSI to their hospital pay-for-reporting program.[12,13] In addition, CMS no longer reimburses for PE/DVTs following total hip or knee replacements.[14]

Given the adverse clinical consequences, as well as associated institutional penalties for pPE/DVT occurrences, we need to understand the extent to which these events are related to prophylaxis-related process-of-care failures and therefore potentially preventable through improved care. Although several recent case series have examined guideline-adherent thromboprophylaxis rates in either at-risk patients or those experiencing pPE/DVT events, data are limited, and somewhat conflicted, on whether thromboprophylaxis use differs between patients who do and

Article-at-a-Glance

Background: Observational studies continue to report thromboprophylaxis underuse for postoperative pulmonary embolism/deep vein thrombosis (pPE/DVT) despite the long-standing existence of prevention guidelines. However, data are limited on whether thromboprophylaxis use differs between patients developing pPE/DVT versus those who do not or on why prophylaxis is withheld.

Methods: Administrative data (2002–2007) from 28 Veterans Health Administration hospitals were screened for discharges with (1) pPE/DVT as flagged by the Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicator software and (2) pharmacoprophylaxis-recommended procedures, and the medical records were reviewed to ascertain true pPE/DVT cases. Controls were selected by matching cases by hospital, age, sex, diagnosis-related group, and predicted probability for developing pPE/DVT, and who underwent a pharmacoprophylaxis-recommended procedure. Records were assessed for "appropriate pharmacoprophylaxis use," defined primarily per American College of Chest Physicians (ACCP) guidelines, and reasons for anticoagulant nonuse.

Results: The 116 case-control pairs were similar in terms of demographics, surgery type, ACCP risk category, and appropriate pharmacoprophylaxis rates overall. Of the highest-risk patients, respective pharmacoprophylaxis rates among cases and controls were 88% versus 92% among hip/knee replacements and 31% versus 48% among cancer patients. Of the cases and controls who did not receive appropriate pharmacoprophylaxis, only about 25% had documented contraindications. Reviewers identified contraindications in 14% of cases and 9% of controls.

Conclusions: Similarities in preventive pPE/DVT practice between cases and controls suggest that pPE/DVTs occur despite implementation of guideline-adherent practices.

those who do not develop pPE/DVT.[7–9] Physician implicit review of chart-confirmed pPE/DVT cases (*N* = 28), using 1994 data, revealed a potential process failure in 61% versus only 2% of controls.[15] However, a subsequent case-control study, using 1995–2004 chart data, reported similar prophylaxis adherence in cases and controls (*N* = 172).[16] In addition, relatively few recent studies have reported on thromboprophylaxis use among high-risk patients,[9,16,17] or on the reasons for or appropriateness of withholding thromboprophylaxis.[18] Further, despite the Veterans Health Administration (VHA)'s strong commitment to development and implementation of quality improvement initiatives, little is known about VTE practices in the VHA.[19]

We therefore undertook this study to better understand the relationship between adherence with guideline-related processes of care and pPE/DVT development and to explore VTE prevention practices in the VHA.

Methods

Study Design and Data Sources

As part of a larger VHA study examining the validity of selected PSIs, from May 2009 through July 2010 we conducted a retrospective case-control study using data from October 1, 2002, through September 30, 2007. We obtained acute care hospital administrative discharge data from the VHA National Patient Care Database Patient Treatment File (PTF),[20,21] and electronic medical record (EMR) data using VistAWeb, a program enabling centralized access to VHA-wide facility data.[22]

Hospital Sampling

Our hospital sampling method is described in detail elsewhere.[23] Briefly, we applied the AHRQ PSI software (v.3.1a) to the PTF–derived database to obtain individual PSI counts and composite scores (a combined measure of 11 PSIs). We selected a geographically diverse sample of 28 of 158 VHA acute care hospitals using a stratified sampling method based on observed and expected PSI counts and composite scores. The final hospital sample included hospitals from 19 of 21 VHA regional health care networks (otherwise known as Veterans Integrated Service Networks), and 20 states representing a mix of rural (for example, Togus, Maine) and more urban areas (for example, Los Angeles). (All the US Census Bureau–designated regions and divisions of the 48 contiguous states and the District of Columbia were represented.)

Pulmonary Embolism/Deep Vein Thrombosis Case and Control Identification

Case and control identification required several steps (Figure 1, page 350). First, as part of a previous study examining the positive predictive value of the PSI pPE/DVT,[23] we randomly selected 112 PSI software–flagged pPE/DVT cases (4 per sample hospital); two trained nurse-abstractors reviewed these EMRs to identify true pPE/DVTs (*N* = 48).[23,24] Second, we excluded cases with ineligible procedures (that is, those not appropriate for prophylaxis [*N* = 25]; *see* below for eligibility); and third, we confirmed there were no upper-extremity DVT cases. This left 23 cases for matching. We then screened additional flagged cases in batches of 112, repeating Step 2 (screening for eligible procedures), then Steps 1 and 3 (nurse-abstractor confirmation of PE/lower-extremity DVT) until we obtained approximately 100 cases. This target sample size was based on power calculations that assumed a pharmacoprophylaxis use difference of 20% between cases and controls and a rate of approximately 50% among cases.[25]

We used the SCIP–recommended procedures for thromboprophylaxis as initial reference to determine surgeries for inclusion.[26] Additional selection was based on study clinician opinion [A.M.B., A.J.C.] with expert surgical input [K.I.]; only major surgeries were included. We excluded neurosurgical procedures since they often do not receive pharmacologic prophylaxis, and included lower-extremity orthopedic procedures associated with fractures (excluding below knee) and lower-extremity amputations lasting beyond one hour.

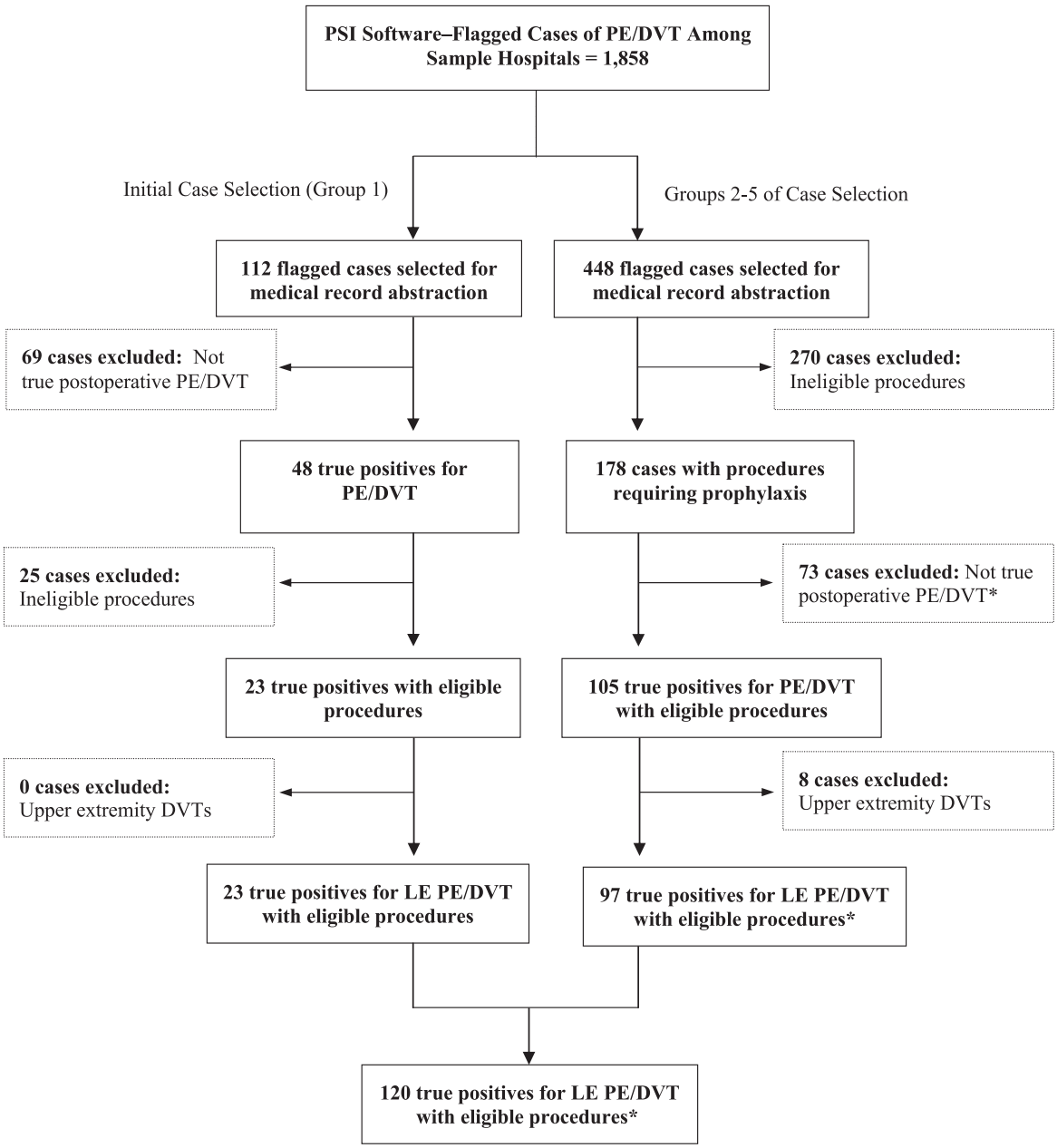
In total, we screened 560 PSI software–flagged cases (five groups of 112, including the initial 112) (Figure 1); among those with eligible procedures, 128 were confirmed as true pPE/DVTs. We excluded 8 upper-extremity DVTs, plus 3 cases occurring more than 30 days after surgery.

Using the PTF, we then matched true pPE/DVT cases one-to-one with controls (that is, patients not experiencing a pPE/DVT), based on hospital, age (within 5 years), sex, and diagnosis-related group ([DRG], which accounts for procedure type). Within strata defined by these variables, we selected the control with the closest predicted probability of developing pPE/DVT based on a logistic regression model that included age, sex, DRG, and discharge comorbidities.[27] We similarly screened out controls on the basis of procedure eligibility as described for cases. We lacked an appropriate match for 1 case; our final sample consisted of 116 matched cases and controls. Our cases came from 27 of the 28 sample hospitals. The average number of cases per site was 4.1 (standard deviation [SD], 2.7) or median 3.5 (range, 0–10).

Electronic Medical Record Abstraction

Starting with preliminary tools and guidelines from AHRQ,[28]

Selection of Pulmonary Embolism (PE)/Deep Venous Thrombosis (DVT) Cases



* From groups 2–5; 3 additional cases were excluded for PE/DVT occurring > 30 days postoperatively, and 1 case was excluded because of no control match, yielding a final sample of 116.

Figure 1. *Case and control identification required several steps, as shown. LE, lower extremity.*

we developed separate but comparable abstraction instruments for cases and controls. Specific to cases, the instrument included questions about case ascertainment per the PSI definition.[24] There had to be documentation in the discharge summary, radiology reports, or physician progress notes that the patient experienced a VTE postoperatively. All but 2 cases had radiographic confirmation; these 2 cases were reviewed by study clinicians [A.M.B., A.J.C.] and retained as true events. (See Kaafarani et al. for further EMR abstraction details.[23])

For both cases and controls, the instrument contained iden-

Table 1. Patient Sample Characteristics*

Characteristic	Cases (*N* = 116)	Controls (*N* = 116)
Age, years, mean (SD)	70.0 (10.2)	70.1 (10.3)
Sex, male *n* (%)	115 (99.1)	115 (99.1)
Length of stay, days, median (25th, 75th percentiles)†‡	18.0 (10.0, 32.0)	9.0 (5.0, 15.0)
Race/Ethnicity *n* (%)		
White, non-Hispanic	83 (71.6)	85 (73.3)
Non-white	19 (16.4)	13 (11.2)
Other/missing	14 (12.1)	18 (15.5)
Patient Risk Factors Present on Admission *n* (%)		
Current neoplasm§	32 (27.6)	29 (25.0)
Obesity (BMI ≥ 30)†‖	25 (21.6)	42 (36.2)
Prior venous thromboembolism†	18 (15.5)	6 (5.2)
History of recent trauma	7 (6.0)	12 (10.3)
Hypercoaguable state#	1 (0.9)	1 (0.9)
ACCP 6/7 Highest-Risk Category**	95 (81.9)	90 (77.6)

* Column percentages are shown. SD, standard deviation; BMI, body mass index; ACCP, American College of Chest Physicians.

† $p < .05$; significant difference between cases and controls.

‡ Median time from admission to venous thromboembolism (VTE) diagnosis was 6.5 days (25th and 75th percentiles: 3.0, 12.0); median time from diagnosis to discharge was 10.0 days (25th and 75th percentiles: 5.0, 19.0).

§ Includes known neoplasms present on admission and masses present on admission that were diagnosed as malignancies.

‖ BMI based on admission height and weight.

Includes congenital and acquired states other than malignancy; 1 case and 1 control had a known Factor V Leiden mutation (and a prior VTE).

** Based on procedure (hip or knee replacement or hip fracture surgery) and/or age > 40 years and having a prior VTE/molecular hypercoagulability or current neoplasm. (Geerts WH, et al. Prevention of venous thromboembolism. *Chest.* 2001;119(1 Suppl):132S–175S [reference 31]; Geerts WH, et al. Prevention of venous thromboembolism: The Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest.* 2004;126(3 Suppl):338S-400S. [reference 32].) Major trauma and acute spinal cord injury patients are also at the highest risk. Only 1 control fit these latter criteria.

tical questions about demographics, patient- and procedure-related risk factors, and perioperative thromboprophylaxis use, plus contraindications to pharmacoprophylaxis, including a history of recent (within four weeks) or active intracerebral, gastrointestinal, respiratory, or urinary tract bleeding, a known bleeding disorder (inherited or acquired, such as severe liver disease), or platelet count < 75,000.[29,30] The two nurse-abstractors were also encouraged to write-in additional documented reasons for lack of pharmacoprophylaxis use. For controls, records were also examined out to 90 days following discharge to look for late VTE events. (We initially tested interrater reliability on approximately 10% of cases; average agreement across records was 92%. Further abstraction details are described elsewhere.[23])

Analyses

We compared cases and controls on several variables, including demographics, patient-related risk factors, surgery type, American College of Chest Physicians (ACCP) risk category (Table 1, above),[31,32] and thromboprophylaxis use. We conducted similar analyses comparing prophylaxis use between and within surgical specialties, with subcategorization of orthopedic procedures, and compared risk factors within surgery type.

Thromboprophylaxis Use. We categorized thromboprophylaxis use into mutually exclusive groups: "appropriate pharmacoprophylaxis," "mechanical prophylaxis only," and "no prophylaxis." We used the SCIP VTE measures specifications (v.2.6b), derived from the seventh version of the ACCP guidelines (ACCP 7), to define appropriateness.[27,32] Appropriate pharmacoprophylaxis required both proper timing of administration (within 24 hours of the procedure) and use of recommended anticoagulants; the "mechanical prophylaxis only" group similarly required use of appropriately timed and recommended forms of mechanical prophylaxis (Appendix 1, available in online article). For procedures not represented in the SCIP measures (for example, lower-extremity amputations), we referred directly to ACCP 7 for recommended anticoagulants.[32] Of note, the SCIP specifications differ from ACCP guidelines in being more explicit with respect to appropriate timing of prophylaxis administration, outlining appropriate prophylaxis options for a specific set of major surgeries, and not accounting for patient VTE risk factors (for example, SCIP considers mechanical prophylaxis as appropriate in all urologic patients, even high-risk ones).[26,32] Because ACCP guidelines recommend that mechanical prophylaxis be used as an adjunct to anticoagulants, particularly in high-risk patients, we

also calculated the percentage on both appropriate pharmacologic and mechanical prophylaxis.[32]

Reasons for Lack of Pharmacoprophylaxis. Study clinicians [A.M.B., A.J.C] reviewed EMRs of all patients who did not receive appropriate pharmacoprophylaxis and in whom nurse-abstractors did not identify a potential contraindication (including records with delayed prophylaxis; that is, initiated > 24 hours postoperatively). They looked for other potential reasons for lack of pharmacoprophylaxis use, such as significant perioperative bleeding (for example, bleeding-associated systolic blood pressure drop to < 90 mm Hg, or hematocrit drop to < 25% immediately postoperatively, or requiring ≥ four units of blood), significant renal impairment (estimated glomerular filtration rate < 30ml/min or end-stage renal disease on dialysis), or epidural anesthesia/analgesia use. This last item was included because the SCIP specifications consider withholding of pharmacoprophylaxis in patients undergoing epidural anesthesia or with an epidural catheter in situ as appropriate, provided mechanical prophylaxis is used, contrary to ACCP guidelines.[1,32]

We compared groups using parametric (chi-squares and *t*-tests) and nonparametric tests (Wilcoxon rank-sum) as appropriate. We also calculated effect size (ES) for selected results to characterize the clinical significance of findings.[33] (Effect sizes of 0.2, 0.5, and 0.8 are considered small, medium, and large respectively.[33]) SAS version 9.1 (SAS Institute Inc., Cary, North Carolina) was used for all analyses.

Results

Patient Sample Characteristics

Table 1 shows baseline characteristics of sample patients. Cases and controls were similar with respect to sex, race/ethnicity, and age; all patients were at least 40 years old. The most common preadmission risk factors overall were current neoplasm and obesity (that is, body mass index [BMI] ≥ 30). Individual risk factors were similar between cases and controls, except for more prior VTEs in cases (16% versus 5%; $p = .01$) and more obesity in controls (36% versus 22%; $p = .03$). Cases and controls were comparable with respect to major surgery type; orthopedic surgery was the most common category (> 54% of cases and controls) followed by general surgery (> 29% of cases and > 26% of controls; Table 2, page 353). Eighty-two percent of cases and 77% of controls were in the ACCP highest risk category ($p = .42$); all were at least high risk.

There were few significant differences in risk factors between cases and controls by surgery type, other than higher prior VTE rates among orthopedic cases and impaired mobility among lower-extremity amputation cases (data not presented; available from authors.) Notably, current neoplasm was particularly common among patients undergoing thoracic (100% of cases, 88% of controls), urologic (88% of cases, 71% of controls), and general surgery (44% of cases, 52% of controls). We found no late VTE events among controls.

Overall Thromboprophylaxis Rates

As shown in Table 2, more controls than cases received appropriate pharmacoprophylaxis (72% versus 62%), although the difference was not statistically significant ($p = .13$; ES = .16). Mechanical prophylaxis alone was used in 32% of cases and 22% of controls ($p = .11$); 6% of both cases and controls did not receive any prophylaxis. Fifty-three percent of cases and 58% of controls received both appropriate pharmacologic and mechanical prophylaxis ($p = .60$).

Surgical Specialty-Specific Thromboprophylaxis Rates

Orthopedic patients were significantly more likely to receive appropriate pharmacoprophylaxis compared with other surgical groups ($p < .01$), with similar rates among cases and controls (> 87%). Urologic patients had the lowest rates (0% and 29% for cases and controls, respectively). There was a nonsignificant trend toward higher prophylaxis rates among controls versus cases for the nonorthopedic surgical groups.

We saw opposite results with respect to mechanical prophylaxis only use. Nonorthopedic surgical patients (other than lower-extremity amputations and thoracic patients) had significantly higher mechanical prophylaxis rates compared with orthopedic patients ($p < .001$); urologic patients had the highest rates (88% of cases and 71% of controls).

Rates of no prophylaxis were comparable between cases and controls by specialty group, ranging from 0% to 50% for cases undergoing orthopedic procedures versus lower-extremity amputations. With respect to use of both pharmacologic and mechanical prophylaxis, trends were similar to those seen with appropriate pharmacoprophylaxis.

ACCP Highest-Risk Groups

Eighty-one percent of both cases and controls undergoing orthopedic procedures represented hip or knee replacements. Among these joint replacement patients, 88% of cases and 92% of controls received appropriate pharmacoprophylaxis ($p = .65$); all received some form of prophylaxis; approximately three quarters of both groups received a combination of appropriate pharmacologic and mechanical prophylaxis (Table 2.)

While 83% of patients with a prior VTE in both groups

Table 2. Rates of Thromboprophylaxis Among All Surgical Patients and by Surgery Type*

	N		Appropriate Pharmacoprophylaxis† *n* (%)		Mechanical Prophylaxis Only‡ *n* (%)		No Prophylaxis *n* (%)		Appropriate Pharmaco- and Mechanical† *n* (%)	
Surgery Type	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
All Procedures	116	116	72 (62.1)	83 (71.6)	37 (31.9)	26 (22.4)	7 (6.0)	7 (6.0)	62 (53.4)	67 (57.8)
Orthopedic§	63	63	55 (87.3)	55 (87.3)	8 (12.7)	7 (11.1)	0	1 (1.6)	49 (77.8)	46 (73.0)
Knee Replacement	31	25	28 (90.3)	23 (92.0)	3 (9.7)	2 (8.0)	0	0	26 (83.9)	20 (80.0)
Hip Replacement	20	26	17 (85.0)	24 (92.3)	3 (15.0)	2 (7.7)	0	0	14 (70.0)	19 (73.1)
General Surgery	34	31	13 (38.2)	17 (54.8)	19 (55.9)	12 (38.7)	2 (6.0)	2 (6.5)	12 (35.3)	16 (51.6)
Urologic	8	7	0	2 (28.6)	7 (87.5)	5 (71.4)	1 (12.5)	0	0	2 (28.6)
Thoracic	5	8	2 (40.0)	5 (62.5)	2 (40.0)	2 (25.0)	1 (20.0)	1 (12.5)	1 (20.0)	3 (37.5)
LE Amputation‖	6	7	2 (33.3)	4 (57.1)	1 (16.7)	0	3 (50.0)	3 (42.9)	0	0

* There were no significant differences between cases and controls for "all procedures" or procedure-specific comparisons. Row percentages are shown. LE, lower extremity.

† Appropriateness definition based on administration timing and recommended medications per the Surgical Care Improvement Project (SCIP) Specifications Manual and American College of Chest Physicians (ACCP) guidelines. (QualityNet. Specifications Manual, Version 2.6b: Discharges 04/01/2009 to 09/30/2009. Accessed Jun 27, 2012. http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1221491528970. Appendix A - ICD-9 Code Tables, revised Nov 19, 2008 [Tables 5.19, 5.21, 5.22, 5.23, and 5.24] [reference 26]; and Geerts WH, et al. Prevention of venous thromboembolism: The Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest.* 2004;126(3 Suppl):338S–400S [reference 32]; see Appendix 1 [available in online article].)

‡ Initiated within 24 hours of the first major operation and adhered to recommended SCIP specifications and ACCP guidelines with respect to mechanical prophylaxis type.

§ Open reduction internal fixations of fractures, including proximal femoral fractures, are not included in the SCIP measure (QualityNet. Specifications Manual, Version 2.6b: Discharges 04/01/2009 to 09/30/2009. Accessed Jun 27, 2012. http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1221491528970. Appendix A - ICD-9 Code Tables, revised Nov 19, 2008 [Tables 5.19, 5.21, 5.22, 5.23, and 5.24] [reference 26]).

‖ Not included in the SCIP VTE measure. Removing these patients produces similar "all procedure" rates (for example, pharmacoprophylaxis rates would be 64% in cases and 72% in controls).

received pharmacoprophylaxis, only 31% of cases and 48% of controls with a neoplasm received appropriate pharmacoprophylaxis (p = .19; ES = 0.33). Of those in the highest ACCP risk category based on either procedure or patient risk factors, 68% of cases and 75% of controls received pharmacoprophylaxis (p = .33); 56% of cases and 62% of controls received both pharmacologic and mechanical prophylaxis (p = .46).

Reasons for Lack of Pharmacoprophylaxis

Nurses identified contraindications in only approximately 25% of both cases and controls who did not receive appropriate pharmacoprophylaxis. Table 3 (page 354) shows results after clinician EMR review. Clinicians identified bleeding contraindications in 12% of all cases (N = 14) and 9% of controls (N = 10). Of these, 57% of cases (N = 8) and 60% of controls (N = 6) had these reasons explicitly documented, and 57% (N = 8) and 70% (N = 7), respectively, represented patients with cancer. Another 2 cases had documentation of not receiving pharmacoprophylaxis because of fall risk. Among patients with contraindications, all but 1 case, a patient with peripheral vascular disease, received appropriate mechanical prophylaxis. Two additional cases developed VTEs within 24 hours postoperatively. We felt that we could not conclude that they lacked a contraindication to pharmacoprophylaxis. Given the relatively short time frame between the procedure and the VTE event, postoperative documentation of the acute complication would likely have taken precedence over documentation of contraindications. (Had we excluded these patients from the original sample, our results would not have appreciably changed.)

Thus, accounting for contraindications and early VTE occurrence, a total of 78% of cases and 80% of controls were appropriately managed (p = .75). This represented 56% of cases and 72% of controls with malignancy (p = .19; ES = .32). Accounting for contraindications, assessments of appropriate management increased for all surgical groups except thoracic cases and lower extremity amputation controls (Figure 2, page 355.)

Of the remaining patients, either no reason was given or identified in the record as to why they did not receive appropriate pharmacoprophylaxis (16% of both cases and controls), or potential reasons for withholding prophylaxis were not consistent with accepted contraindications (that is, the presence of an epidural catheter or use of aspirin as prophylaxis instead of anticoagulants; 6% of cases and 4% of controls). No identifiable cause for lack of pharmacoprophylaxis use was found proportionately most often in lower-extremity amputation patients. Of note, among patients with "delayed" prophylaxis (that is, started

Table 3. Reason for Inappropriate or No Pharmacoprophylaxis*

	N		Bleeding† *N* (%)		No Cause Given/Identified *n* (%)		Epidural in Place‡ *n* (%)		Inappropriate Prophylaxis Type§ *n* (%)		Other‖ *n* (%)	
Surgery Type	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
All Procedures	116	116	14 (12.1)	10 (8.6)	19 (16.4)	18 (15.5)	6 (5.2)	4 (3.4)	1 (0.9)	1 (0.9)	4 (3.4)	0
Orthopedic	63	63	0	2 (3.2)	6 (9.5)	5 (7.9)	0	0	1 (1.6)	1 (1.6)	1 (1.6)	0
General Surgery	34	31	9 (26.5)	4 (12.9)	8 (23.5)	8 (25.8)	2 (5.9)	2 (6.5)	0	0	2 (5.9)	0
Urologic	8	7	4 (50.0)	3 (42.9)	1 (12.5)	2 (28.6)	2 (25.0)	0	0	0	1 (12.5)	0
Thoracic	5	8	0	1 (12.5)	1(20.0)	0	2 (40.0)	2 (25.0)	0	0	0	0
LE Amputation	6	7	1 (16.7)	0	3 (50.0)	3 (42.9)	0	0	0	0	0	0

* There were no significant differences between cases and controls overall. Row percentages are shown. LE, lower extremity.

† Including a history of bleeding on admission, intraoperatively, or postoperatively, or a bleeding history with anticoagulation (1 case and 1 control), or at bleeding risk because of thrombocytopenia (platelets < 75, 1 case) or end-stage renal disease (1 case).

‡ These were epidural catheters that stayed in place for ≥ 1 day postoperatively for analgesia; 1 case had this explicitly noted by a clinician as a reason for withholding anticoagulation. An additional 3 cases (2 orthopedic and 1 LE amputation) and 5 controls (3 orthopedic, 1 general surgery, and 1 LE amputation) had an epidural as anesthesia for surgery only.

§ One case and 1 control, each of whom underwent elective knee replacements, received aspirin as prophylaxis.

‖ Includes 2 cases with a pulmonary embolism within 1 day postoperatively who did not receive or have any pharmacoprophylaxis ordered before the event, and 2 cases—including 1 patient undergoing a hip open reduction internal fixation—who were deemed to be at too high fall risk.

in hospital but > 24 hours postoperatively; 2 cases and 8 controls) we did not find documentation of resolution of a contraindication (accepted or otherwise), except for 2 controls in whom it was started after an epidural catheter was removed.

Discussion

This study is one of the few investigations of VHA surgical patients to examine thromboprophylaxis use and the first VHA study to our knowledge to evaluate the association of process failures with outcomes. Among this high-risk group of veterans undergoing major operative procedures, we did not find a significant difference in pharmacoprophylaxis use between cases and controls. Overall, appropriate pharmacoprophylaxis rates were modest, although rates of appropriate management were high when accounting for contraindications; 62% of cases and 72% of controls received guideline-recommended PE/DVT pharmacoprophylaxis, with a further 16% of cases and 8% of controls appropriately managed given contraindications. Had we approximated SCIP criteria by considering mechanical prophylaxis use as appropriate in all urologic patients or patients receiving epidural anesthesia[26,32] and by excluding non-SCIP-eligible procedures (that is, lower-extremity amputations and fractures), 83% of our cases and 90% of controls would have received appropriate prophylaxis (not accounting for SCIP exclusion criteria).[21,27]

Pharmacoprophylaxis rates varied widely by surgery type, with the highest rates among orthopedic patients; general surgery, thoracic, and urologic patients had comparatively lower rates. Among the highest ACCP risk groups, very high pharmacoprophylaxis rates were seen among joint replacement patients and those with a VTE history (> 80%), with much lower rates among cancer patients (39% overall). These lower rates persisted even after accounting for contraindications. (Although the clinical significance—that is, effect size—of case-control differences was slightly larger among cancer patients, both groups were relatively undertreated.)

Clinical trial results strongly support the process-outcome link between perioperative PE/DVT pharmacoprophylaxis use and decreased occurrence of DVT, PE, and death, with much weaker evidence supporting mechanical prophylaxis and only with respect to DVT risk reduction.[1] Thus, the ACCP guidelines generally recommend mechanical prophylaxis alone only when the bleeding risk is unacceptably high.[1] Although we did find a trend toward lower pharmacoprophylaxis use (and conversely higher mechanical prophylaxis use) in cases compared to controls (both overall and within each nonorthopedic surgical group), these differences were not statistically significant. Interestingly, when we performed a post hoc analysis among the nonorthopedic patients, we found borderline significant differences in pharmacoprophylaxis rates (32% in cases versus 53% in controls; $p = .049$, ES = 0.39). However, after we accounted for contraindications, differences in appropriate prophylaxis use disappeared (64% versus 66%; $p = 1.0$, ES = .03). Thus, although

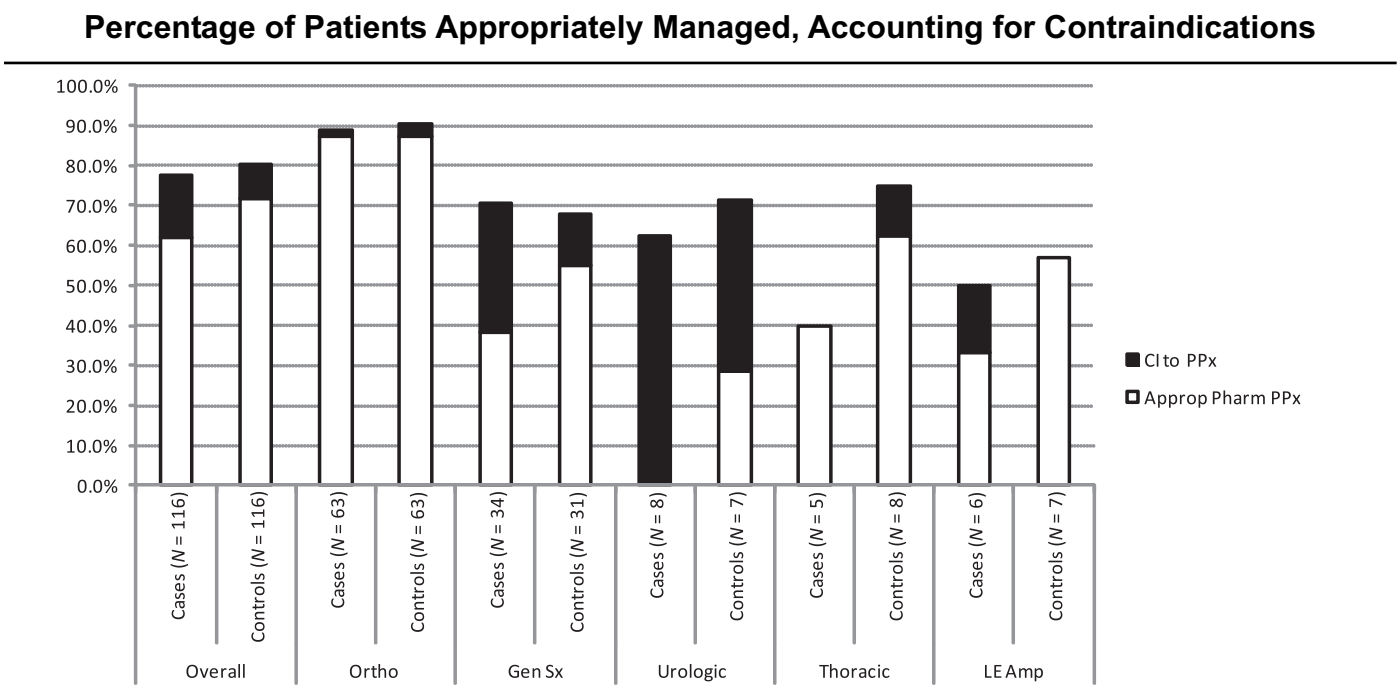


Figure 2. *Accounting for contraindications to pharmacoprophylaxis, assessments of appropriate prophylaxis increased overall and for all surgical groups, except thoracic cases and lower-extremity amputation controls. For example, overall, 62% of cases received appropriate pharmacoprophylaxis; an additional 16% either had accepted contraindications to this but received appropriate mechanical prophylaxis or had early occurrence of venous thromboembolism. If this subgroup is considered as appropriately managed, then the percentage of cases appropriately managed increased from 62% to 78%. CI to PPx, contraindication to prophylaxis; Approp Pharm PPx , appropriate pharmacologic prophylaxis; Ortho, orthopedic; Gen Sx, general surgery; LE Amp, lower-extremity amputation.*

more cases did not receive pharmacoprophylaxis, which may have contributed to at least some events, several of these were appropriately treated and therefore not clearly preventable. Moreover, these findings remain consistent with existing observational studies, both case-control and case series, that show that many pPE/DVT events occur despite the apparent use of best practices.[18,25,34]

A similar disconnect between observational study and clinical trial results with respect to process-outcome links has been noted for various conditions.[35–37] Clinical trials occur in a structured setting with a relatively homogeneous population; few interventions are as effective in actual practice. Further, as hypothesized by Shackford et al., guideline-recommended measures may simply be "insufficient to prevent VTE in some high-risk patients."[16(p. 9)] However, we are not advocating for a decrease in prophylaxis use. Rather, our findings suggest opportunities for improvements in preventive care, particularly among cancer patients and those undergoing nonorthopedic procedures. Although each surgical specialty has its own guidelines, the ACCP VTE prophylaxis guidelines are considered the standard of care in the United States. Clinicians during our study period (2002–2007) would have been influenced by both the sixth (released in 2001) and seventh (released in 2004) versions of these guidelines, which contained very similar recommendations.[31,32] Orthopedic procedures have the highest procedure-associated VTE risk, the strongest evidence base for prevention, and, therefore, the strongest recommendations for pharmacoprophylaxis. However, ACCP 6 and 7 also strongly recommend pharmacoprophylaxis for patients undergoing major general and thoracic procedures, particularly if they are older than 40 years of age or have other VTE risk factors; patients older than 40 years of age with a prior VTE or cancer are considered at highest risk, similar to orthopedic patients.[31,32] For major urologic or vascular procedures, the guidelines recommend anticoagulant use if patients have VTE risk factors. Thus, several of our high- and highest-risk cases and controls did not receive guideline-recommended care (even accounting for contraindications). Of note, our findings would be similar even accounting for the recommendations of the recently published ACCP 9, other than one notable exception; aspirin is now considered an acceptable anticoagulant alternative in patients undergoing major orthopedic surgery.[38] Among patients undergoing knee arthroplasty, we had one case and one control who received only aspirin.

Nevertheless, our rates of appropriate pharmacoprophylaxis

are generally higher than those reported in recent case series of surgical patients.[7,9,16,18,39,40] A multinational trial using chart review (the ENDORSE study) and two large studies in the United States that used administrative data from similar time periods reported overall appropriate pharmacoprophylaxis rates between 18% and 55%.[7,9,39,40] Rates also varied by specialty, with the highest rates among orthopedic patients (38% to 74%), and the lowest rates tending to be among urologic patients (26% to 32%).[7,9,40] The ENDORSE study also found that patients with prior VTE had a higher likelihood of receiving recommended prophylaxis than cancer patients.[9] Among studies of surgical patients experiencing VTE, two single site studies, one using 2009 data (N = 89), and the other, the previously noted case-control study, found that 63% and 40%, respectively, received appropriate pharmacoprophylaxis; an additional 25% and 23%, respectively were appropriately managed given contraindications.[18]

There are several possible explanations for the discrepant rates, particularly with the larger case series, including slightly different appropriateness definitions (all were based on ACCP 6 or 7 guidelines),[31,32] data sources used, patient characteristics, and included procedure types. Our population's higher VTE risk likely led to higher prophylaxis rates. For example, in the ENDORSE study, rates of prior VTE, active cancer, and orthopedic procedures were only 3%, 17%, and 13%, respectively.[9]

Other important findings of our study were that relatively few of the patients who did not receive pharmacoprophylaxis had documented contraindications and that approximately half of these cases and controls had no identifiable reason for lack of use even after careful clinician review. Few previous studies have investigated this aspect of prophylaxis, other than reporting bleeding risk.[9,39] In a setting with an established VTE prophylaxis protocol, including an order sheet with check boxes to indicate contraindications, Weigelt et al. found that only 1% of cases were missing documentation of contraindications.[18] However, Shackford et al. found up to 37% of VTE cases had no contraindications documented or identified.[16]

Our study design is unique compared to related recent studies in surgical patients in that we used a well-established administrative data–based indicator, the AHRQ PE/DVT PSI, to flag PE/DVT cases.[16,18,34] However, given the PSI's recognized relatively high rate of misidentification of events (and underdetection of events), we then confirmed cases through chart review.[23,41,42]

Limitations

Our study had a few limitations. First, our sample size may have been too small to show statistical significance. However, the calculated effect size for the level of difference observed for the whole sample was relatively small at 0.16.[33] Effect size is independent of the sample size. Thus, even if we had found a statistically significant result by including twice as many cases and controls, this difference is too small to be clinically meaningful. In addition, as noted, other studies support this lack of difference in management of cases and controls.[16] Second, although we matched cases and controls on VTE risk using administrative data, chart-abstracted data demonstrated some individual risk factor differences between groups, notably with respect to prior VTE and obesity. However, by chart data all patients were in the high or highest ACCP VTE risk categories, with similar distributions by group. Third, despite using very accurate medication data (obtained from bar-coded medication administration logs), we may be overestimating appropriate pharmacoprophylaxis rates, given that we assessed prophylaxis use up to three days postoperatively but not beyond this. Further, although standard prophylaxis dosing for unfractionated and low molecular weight heparin had to be adhered to, target International Normalized Ratios (INRs) were frequently not documented. Because we could not determine warfarin-dosing appropriateness, we deemed any regular warfarin use as appropriate. Nevertheless, our rates are still higher than those in the ENDORSE study, which did not consider dosage and only examined prophylaxis use at a single time-point. In addition, our methods of assessing appropriateness are similar to those of other recent chart-review studies.[16,18,39] Finally, it is possible that some controls may have had subclinical VTE. However, our findings are consistent with related studies that included only symptomatic VTEs.[16,18] Although we focused on in-hospital outcomes, related studies used 30-day outcomes.[16,18] More recent data have also raised concern about events occurring up to 90 days following surgery, particularly hip replacements.[43] We therefore examined EMRs of all controls up to 90 days postdischarge and found no VTE events; all had evidence of ongoing VHA use after discharge to at least 90 days, unless they died in-hospital or prior to day 90. (None of the deaths were attributed to a VTE.)

Implications and Next Steps

Our findings and those of others suggest that even with 100% compliance with current SCIP VTE measures or even with modifications to increase use of pharmacoprophylaxis, pPE/DVTs will still occur. It may be important to incorporate other factors such as VTE risk or anticoagulant dosage and duration into VTE process measures to realize the full potential of such measures in improving patient safety. Moreover, penalizing providers for pPE/DVT events, such as events that follow joint replacements,

may be inappropriate because the vast majority of pPE/DVT events occur despite appropriate pharmacoprophylaxis. In addition, our finding that clinicians often failed to document reasons for lack of pharmacoprophylaxis is concerning from a medicolegal standpoint. We expect that such documentation will improve over time as provider awareness of performance monitoring and public reporting increases because this information is necessary to satisfy the SCIP VTE measures for certain surgery types.

Further work is needed to understand why these pPE/DVT events are happening. This may occur at the local level by means of observational studies, including local quality improvement activities that carefully assess process factors, using prospective data collection to determine whether important strategies such as early mobilization are actually being implemented. Clinical trials are also necessary to examine additional strategies to reduce VTE events.

Conclusion

Similarities in pPE/DVT preventive practices between cases and controls suggest that pPE/DVT events occur even with implementation of evidence-based practices. However, despite high overall rates of guideline-adherent care, certain high-risk patient groups were relatively undertreated, and documentation of prophylaxis contraindications was frequently absent. While these deficiencies should be addressed through VHA quality improvement efforts, further research is necessary to uncover additional methods to prevent pPE/DVT. J

Funding was provided by the US Department of Veterans Affairs Health Services Research and Development Service grant number SDR 07-002. The authors extend special thanks for their assistance to Kathleen Hickson, RN, PhD, and Sally MacDonald, RN, for abstraction of electronic medical records and modification of the abstraction instrument; Haytham Kaafarani, MD, MPH, for development of the abstraction instrument and case review; Amresh Hanchate, PhD, for development and analysis of predicted probability scores and matching methods; Elaine Hickey, RN, for assessment of interrater reliability; Hillary Mull, MPP, and Qi Chen, MD, MPH, for development of the electronic abstraction form; Susan Loveland, MAT, for programming expertise and identification of the hospital sample, flagged cases, and matched controls; Qi Chen and William O'Brien, MS, for programming; and Elaine Hylek, MD, MPH, for manuscript review and comments.

Online-Only Content

See the online version of this article for

Appendix 1. Surgical Care Improvement Project (SCIP) Venous Thromboembolism (VTE) Measures (v2.6, revised Jan. 9, 2009) and Recommended Prophylaxis

Ann M. Borzecki, MD, MPH, is Research Scientist, US Department of Veterans Affairs (VA), Center for Health Quality, Outcomes and Economic Research, Bedford VA Hospital, Bedford, Massachusetts; and Research Associate Professor, Schools of Public Health and Medicine, Boston University. **Andrew J. Cowan, MD,** is Clinical Instructor in Medicine, Amyloid Treatment and Research Program, Boston Medical Center; and Boston University School of Medicine. **Marisa Cevasco, MD, MPH,** is Research Fellow, Department of Surgery, Harvard University School of Medicine; and Brigham and Women's Hospital, Boston. **Marlena H. Shin, JD, MPH,** is Research Health Scientist, Center for Organization, Leadership and Management Research, VA Boston Healthcare System. **Michael Shwartz, PhD,** is Senior Research Scientist, Center for Organization, Leadership and Management Research, VA Boston Healthcare System; and Richard D. Cohen Professor in Management, School of Management, Boston University. **Kamal Itani, MD,** is Chief of Surgery, VA Boston Healthcare System; and Professor of Surgery, Departments of Surgery, Harvard University and Boston University Schools of Medicine. **Amy K. Rosen, PhD,** is Senior Research Scientist Center for Organization, Leadership and Management Research, VA Boston Healthcare System; and Professor, Schools of Public Health and Medicine, Boston University. Please address correspondence and requests for reprints to Ann M. Borzecki, amb@bu.edu.

References

1. Geerts WH, et al. Prevention of venous thromboembolism: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest.* 2008;133(6 Suppl):381S–453S.
2. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA.* 2003 Oct. 8;290(14):1868–1874.
3. Gangireddy C, et al. Risk factors and clinical impact of postoperative symptomatic venous thromboembolism. *J Vasc Surg.* 2007;45(2):335–341; discussion 341–342.
4. Collins R, et al. Reduction in fatal pulmonary embolism and venous thrombosis by perioperative administration of subcutaneous heparin. Overview of results of randomized trials in general, orthopedic, and urologic surgery. *N Engl J Med.* 1988 May 5;318(18):1162–1173.
5. Mismetti P, et al. Meta-analysis of low molecular weight heparin in the prevention of venous thromboembolism in general surgery. *Br J Surg.* 2001;88(7):913–930.
6. Hull RD, Raskob GE, Hirsh J. Prophylaxis of venous thromboembolism. An overview. *Chest.* 1986;89(5 Suppl):374S–383S.
7. Amin A, et al. Are hospitals delivering appropriate VTE prevention? The venous thromboembolism study to assess the rate of thromboprophylaxis (VTE start). *J Thromb Thrombolysis.* 2010;29(3):326–339.
8. Dorfman M, Chan SB, Maslowski C. Hospital-acquired venous thromboembolism and prophylaxis in an integrated hospital delivery system. *J Clin Pharm Ther.* 2006;31(5):455–459.
9. Kakkar AK, et al. Venous thromboembolism risk and prophylaxis in the acute care hospital setting (ENDORSE survey): Findings in surgical patients. *Ann Surg.* 2010;251(2):330–338.
10. The Joint Commission: Specifications Manual for National Hospital Inpatient Quality Measures. Version 4.1.1. Accessed Jun 28, 2012. http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx.
11. Agency for Healthcare Research and Quality. AHRQ Quality Indicators: Guide to Patient Safety Indicators. Version 3.1. Mar 12, 2007. Accessed Jun 28, 2012. http://qualityindicators.ahrq.gov/Downloads/Software/SAS/V31/psi_guide_v31.pdf.

12. Centers for Medicare & Medicaid Services. Fiscal Year 2009 Quality Measure Reporting for 2010 Payment Update. Accessed Jun 28, 2012. http://www.cms.hhs.gov/HospitalQualityInits/downloads/HospitalRHQDAPU200808.pdf.
13. QualityNet. Hospital Inpatient Quality Reporting (IQR) Program Overview: Resources for Fiscal Year 2013 Payment Determination. Program Changes/Updates. Accessed Jun 28, 2012. http://www.qualitynet.org/dcs/ContentServer?cid=1138115987129&pagename=QnetPublic%2FPage%2FQnetTier2&c=Page.
14. Centers for Medicare & Medicaid Services. Hospital-Acquired Conditions (Present on Admission Indicator). (Updated May 1, 2012.) Accessed Jun 28, 2012. http://www.cms.gov/HospitalAcqCond/06_Hospital-Acquired_Conditions.asp#TopOf Page.
15. Weingart SN, et al. Use of administrative data to find substandard care: Validation of the complications screening program. *Med Care.* 2000;38(8):796–806.
16. Shackford SR, et al. A 10-year analysis of venous thromboembolism on the surgical service: The effect of practice guidelines for prophylaxis. *Surgery.* 2008;144(1):3–11.
17. Schleyer AM, et al. Adherence to guideline-directed venous thromboembolism prophylaxis among medical and surgical inpatients at 33 academic medical centers in the United States. *Am J Med Qual.* 2011;26(3):174–180.
18. Weigelt JA, Lal A, Riska R. Venous thromboembolism prophylaxis in surgical patients: Identifying a patient group to maximize performance improvement. *Jt Comm J Qual Patient Saf.* 2011;37(4):178–183.
19. Herbers J, Zarter S. Prevention of venous thromboembolism in Department of Veterans Affairs hospitals. *J Hosp Med.* 2010;5(1):E21–25.
20. US Department of Veterans Affairs Information Resource Center. VIReC Research User Guide: FY2002 VHA Medical SAS® Inpatient Datasets. Jan 2003. Sep 2007. Accessed Jul 5, 2012. http://www.virec.research.va.gov/RUGs/MedSAS-Archive/IP/RUG-MedSAS-IP-FY06-ER.pdf.
21. Rosen AK, et al. Evaluating the patient safety indicators: How well do they perform on Veterans Health Administration data? *Med Care.* 2005;43(9):873–884.
22. US Department of Veterans Affairs, Veterans Health Administration. VistAWeb Technical Manual. Jan 2010. Accessed Jun 28, 2012. http://www.va.gov/vdl/documents/Clinical/VistaWeb/webv_1_19_tm.pdf.
23. Kaafarani HM, et al. Validity of selected Patient Safety Indicators: Opportunities and concerns. *J Am Coll Surg.* 2011;212(6):924–934.
24. Agency for Healthcare Research and Quality. AHRQ Quality Indicators: Patient Safety Indicators: Technical Specifications. Version 3.1 Mar 12, 2007. Accessed Jun 28, 2012. http://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V31/psi_technical_specs_v31.pdf.
25. Goldhaber SZ, Dunn K, MacDougall RC. New onset of venous thromboembolism among hospitalized patients at Brigham and Women's Hospital is caused more often by prophylaxis failure than by withholding treatment. *Chest.* 2000;118(6):1680–1684.
26. QualityNet. Specifications Manual, Version 2.6b: Discharges 04/01/2009 to 09/30/2009. Accessed Jun 27, 2012. http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1221491528970. Appendix A - ICD-9 Code Tables, revised Nov 19, 2008 (Tables 5.19, 5.21, 5.22, 5.23, and 5.24).
27. Healthcare Cost and Utilization Project. Comorbidity Software, Version 3.7: Overview of Cormorbidity Software. Apr 2012. (Updated Apr 20, 2012.) Accessed Jun 28, 2012. http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp#overview.
28. Agency for Healthcare Research and Quality. AHRQ Quality Indicators Validation Pilot. Accessed Jun 28, 2012. http://qualityindicators.ahrq.gov/ValidationPilot.aspx.
29. National Institute for Health and Clinical Excellence. Venous Thromboembolism: Reducing the Risk. NICE Clinical Guideline 92. Jun 2010. Accessed Jun 28, 2012. http://guidance.nice.org.uk/CG92/NICEGuidance/pdf/English.
30. Baglin T, et al. Guidelines on the use and monitoring of heparin. *Br J Haematol.* 2006;133(1):19–34.
31. Geerts WH, et al. Prevention of venous thromboembolism. *Chest.* 2001;119(1 Suppl):132S–175S.
32. Geerts WH, et al. Prevention of venous thromboembolism: The Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest.* 2004;126(3 Suppl):338S–400S.
33. Cohen J. *Statistical Power Analysis for the Behavioral Sciences,* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
34. Agnelli G, et al. A clinical outcome-based prospective study on venous thromboembolism after cancer surgery: The @RISTOS project. *Ann Surg.* 2006;243(1):89–95.
35. Fonarow GC, et al. Association between performance measures and clinical outcomes for patients hospitalized with heart failure. *JAMA.* 2007 Jan 3;297(1):61–70.
36. Werner RM, Bradlow ET. Relationship between Medicare's Hospital Compare performance measures and mortality rates. *JAMA.* 2006 Dec 13;296(22):2694–2702.
37. Ingraham AM, et al. Association of Surgical Care Improvement Project infection-related process measure compliance with risk-adjusted outcomes: implications for quality measurement. *J Am Coll Surg.* 2010;211(6):705–714.
38. Guyatt GH, et al. Executive summary: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest.* 2012;141(2 Suppl):7S–47S.
39. Cohen AT, et al. Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): A multinational cross-sectional study. *Lancet.* 2008 Feb 2;371(9610):387–394.
40. Amin AN, et al. Preventing venous thromboembolism in US hospitals: Are surgical patients receiving appropriate prophylaxis? *Thromb Haemost.* 2008;99(4):796–797.
41. Rosen AK, et al. Validating the Patient Safety Indicators in the Veterans Health Administration: Do they accurately identify true safety events? *Med Care.* 2012;50(1):74–85.
42. Classen DC, et al. "Global trigger tool" shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood).* 2011;30(4):581–589.
43. Heit JA. Estimating the incidence of symptomatic postoperative venous thromboembolism: The importance of perspective. *JAMA.* 2012 Jan 18;307(3):306–307.

Online-Only Content

Appendix 1. Surgical Care Improvement Project (SCIP) Venous Thromboembolism (VTE) Measures (v2.6, revised January 9, 2009) and Recommended Prophylaxis

Source: QualityNet. Specifications Manual, Version 2.6b: Discharges 04/01/2009 to 09/30/2009. Accessed Jun 27, 2012. http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1221491528970. Appendix A - ICD-9 Code Tables, revised Nov. 19, 2008 (Tables 5.19, 5.21, 5.22, 5.23, and 5.24).

Performance Measure Name: Surgery Patient with Recommended Venous Thromboembolism Prophylaxis Ordered.

Description: Surgery patients with recommended venous thromboembolism (VTE) prophylaxis ordered anytime from hospital arrival to 24 hours after surgery end time.

SCIP VTE-2

Performance measure name: Surgery Patients Who Received Appropriate Venous Thromboembolism Prophylaxis Within 24 Hours Prior to Surgery to 24 Hours After Surgery

Description: Surgery patients who received appropriate venous thromboembolism (VTE) prophylaxis within 24 hours prior to Surgical Incision Time to 24 hours after surgery end time.

SCIP–Recommended VTE Prophylaxis Selection for Surgery

Surgery Type	Recommended Prophylaxis Options
General Surgery (Table 5.19)*	Any of the following: • Low-dose unfractionated heparin (LDUH) • Low molecular weight heparin (LMWH) • Factor Xa Inhibitor (Fondaparinux) • LDUH or LMWH or Factor Xa Inhibitor combined with IPC or GCS
General Surgery with High Risk for Bleeding† (Table 5.19)*	Any of the following: • Intermittent pneumatic compression (IPC) devices • Graduated compression stockings (GCS)
Urologic Surgery (Table 5.21)*	Any of the following: • Low-dose unfractionated heparin (LDUH) • Low molecular weight heparin (LMWH) • Factor Xa Inhibitor (Fondaparinux) • Intermittent pneumatic compression (IPC) devices • Graduated compression stockings (GCS) • LDUH or LMWH combined with IPC or GCS
Elective Total Hip Replacement (Table 5.22)*	Any of the following started within 24 hours of surgery: • Low molecular weight heparin (LMWH) • Factor Xa Inhibitor (Fondaparinux) • Warfarin
Elective Total Knee Replacement (Table 5.23)*	Any of the following: • Low molecular weight heparin (LMWH) • Factor Xa Inhibitor (Fondaparinux) • Warfarin • Intermittent pneumatic compression (IPC) devices • Venous foot pump

(continued on page AP2)

Online-Only Content

Appendix 1. Surgical Care Improvement Project (SCIP) Venous Thromboembolism (VTE) Measures (v2.6, revised January 9, 2009) and Recommended Prophylaxis (continued)

SCIP–Recommended VTE Prophylaxis Selection for Surgery (continued)

Surgery Type	Recommended Prophylaxis Options
Hip Fracture Surgery (Table 5.24)*	Any of the following: • Low-dose unfractionated heparin (LDUH) • Low molecular weight heparin (LMWH) • Factor Xa Inhibitor (Fondaparinux) • Warfarin
Elective Total Hip Replacement with High Risk for Bleeding† (Table 5.22)*	Any of the following: • Intermittent pneumatic compression (IPC) devices • Venous foot pump
Hip Fracture Surgery with High Risk for Bleeding† (Table 5.24)*	Any of the following: • Graduated compression stockings (GCS) • Intermittent pneumatic compression (IPC) devices • Venous foot pump

* See QualityNet. Specifications Manual, Version 2.6b: Discharges 04/01/2009 to 09/30/2009. Accessed Jun 27, 2012. http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1221491528970 (reference 26) for tables of eligible procedures.

† Patients who receive neuraxial anesthesia or have a documented bleeding risk may pass the performance measure if appropriate mechanical prophylaxis is ordered.

AIDS PATIENT CARE and STDs
Volume 26, Number 5, 2012

DOI: 10.1089/apc.2011.0218

Correlates of Antiretroviral and Antidepressant Adherence Among Depressed HIV-Infected Patients

Kathryn A. Bottonari, Ph.D.,[1] Shanti P. Tripathi, M.S.,[2,3] John C. Fortney, Ph.D.,[2,3] Geoff Curran, Ph.D.,[2,3] David Rimland, M.D.,[4] Maria Rodriguez-Barradas, M.D.,[5] Allen L. Gifford, M.D.,[6,7] and Jeffrey M. Pyne, M.D.[2,3]

Abstract

Although crucial for efficacy of pharmacotherapy, adherence to prescribed medication regimens for both antiretrovirals and antidepressants is often suboptimal. As many depressed HIV-infected individuals are prescribed both antiretrovirals and antidepressants, it is important to know whether correlates of nonadherence are similar or different across type of regimen. The HIV Translating Initiatives for Depression into Effective Solutions (HI-TIDES) study was a single-blinded, longitudinal, randomized controlled effectiveness trial comparing collaborative care to usual depression care at three Veterans Affairs HIV clinics. The current investigation utilized self-report baseline interview and chart-abstracted data. Participants were 225 depressed HIV-infected patients who were prescribed an antidepressant ($n=146$), an antiretroviral ($n=192$), or both ($n=113$). Treatment adherence over the last 4 days was dichotomized as "less than 90% adherence" or "90% or greater adherence." After identifying potential correlates of nonadherence, we used a seemingly unrelated regression (SUR) bivariate probit model, in which the probability of adherence to HIV medications and the probability of adherence to antidepressant medications are modeled jointly. Results indicated that 75.5% ($n=146$) of those prescribed antiretrovirals reported 90%-plus adherence to their antiretroviral prescription and 76.7% ($n=112$) of those prescribed antidepressants reported 90%-plus adherence to their antidepressant prescription, while 67% of those prescribed both ($n=113$) reported more than 90% adherence to both regimens. SUR results indicated that education, age, and HIV symptom severity were significant correlates of antiretroviral medication adherence while gender and generalized anxiety disorder diagnosis were significant correlates of adherence to antidepressant medications. In addition, antiretroviral adherence did not predict antidepressant adherence ($\beta=1.62$, $p=0.17$), however, antidepressant adherence did predict antiretroviral adherence ($\beta=2.30$, $p<0.05$).

Introduction

Despite ongoing advances in pharmacotherapy in the treatments of depression and HIV, many depressed HIV-infected individuals experience only partial response due to nonadherence to prescribed regimens. Indeed, depressive symptoms have been repeatedly identified as a consistent, yet mutable, barrier to medical treatment adherence.[1,2] Moreover, depression has been identified as a common barrier to HIV treatment adherence.[3] Yet, despite advances in antidepressant treatment, there is evidence that depressed patients (both HIV-infected and non-HIV–infected) have trouble adhering to prescribed antidepressant regimens.[4,5] Although past research has explored demographic, behavioral, and health-related barriers to adherence to depression and HIV treatment separately, limited research has examined whether common or distinct factors are barriers for individuals being prescribed both types of medications.

[1]Ralph H. Johnson VA Medical Center, Charleston, South Carolina.
[2]Central Arkansas Veterans Healthcare System, North Little Rock, Arkansas.
[3]Division of Health Services Research, UAMS Psychiatric Research Institute,College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas.
[4]VA Medical Center and Emory University School of Medicine, Atlanta, Georgia.
[5]VA Medical Center and Baylor College of Medicine, Houston, Texas.
[6]Center for Health Quality, Outcomes, and Economic Research (CHQOER), HSRD Center of Excellence, Edith Nourse VA Medical Center, Bedford, Massachusetts.
[7]Boston University Schools of Public Health and Medicine, Boston, Massachusetts.

Examination of such variables will likely add to providers' knowledge of treating the "whole" patient, i.e., examination of psychosocial and medical factors that exist for HIV-infected individuals.[6] The present investigation focuses on barriers to adherence for both antidepressant and antiretroviral medications among depressed HIV-infected individuals prior to their enrollment in an intervention trial aimed at improving depression treatment and outcomes among treatment-seeking HIV-infected individuals.

To date, the literature has examined barriers of adherence to antiretrovirals and antidepressants separately. Extensive research has been conducted on identifying barriers to HIV treatment adherence given the virologic consequences of non-adherence (i.e., virologic failure and development of resistance). The most common factors found to impact HIV treatment adherence include complexity of treatment regimens, side effect profiles, excessive substance use, psychopathology (e.g., depression and anxiety), beliefs about medication, self-efficacy, social support, coping style, and memory lapse.[7] In contrast to HIV treatment adherence, research on antidepressant adherence is relatively sparse. Identified barriers to antidepressant adherence have included younger age, female gender, first episode versus recurrent depression, low self-efficacy, lower levels of formal education, severity of symptoms, side effect profiles, beliefs about medication-taking behavior, regimen characteristics, and forgetting.[8–12] Finally, given evidence that many individuals require maintenance treatment beyond remission of depressive symptoms,[13] it is of concern that several investigators have identified premature discontinuation of treatment as a more severe form of nonadherence. One retrospective chart review indicated that only 44% of their sample completed 6 months of treatment, suggesting that the majority of patients do not receive the full benefit of the antidepressants.[14] Given the illness profiles of both depression and HIV, research is needed to identify barriers to adherence for both antidepressant and antiretroviral medications.

More recently, research has begun to emerge examining adherence patterns among depressed HIV-infected patients. Several investigations have demonstrated improved HIV treatment adherence outcomes with use of antidepressant treatment. First, there was preliminary data from a retrospective chart review suggesting that adherence to antiretrovirals was positively correlated with antidepressants adherence ($r = 0.31$; $p < 0.02$) and of improved virologic responses given evidence of adherence to antidepressants (>80%).[15] Second, several investigations have provided evidence of improved adherence to antiretrovirals among depressed individuals taking antidepressants.[5,16,17] Walkup and colleagues[17] reported that prescription of an antidepressant in a prior month increased the odds of adherence to antiretroviral in the subsequent month. Furthermore, Horzberg and colleagues[5] attempted to unpack the relationship between depression, adherence to antiretrovirals, and HIV treatment adherence. They reported that depressed HIV-infected individuals who were highly adherent (>90%) to their antidepressants had similar antiretroviral adherence profiles to nondepressed individuals and that both of these groups had significantly better antiretroviral adherence than those who were both depressed and nonadherent to their antidepressants. Finally, there is evidence that antidepressant treatment also helps improve adherence to complex antiretroviral regimens.[18] However, given evidence that antidepressant and antiretroviral adherence are related, it would be of interest to identify factors that predict adherence to both types of medication.

This investigation is part of the larger, three-site, two-arm, single-blinded, randomized controlled effectiveness trial, HIV Translating Initiatives for Depression into Effective Solutions (HI-TIDES).[19] The goal of the HI-TIDES trial was to compare a collaborative care intervention[20] to treatment as usual (TAU) in the treatment of depression for depressed HIV clinic patients. The larger trial involved an intervention utilizing an offsite HIV depression care team (registered nurse depression care manager, pharmacist, and psychiatrist) who delivered up to 12 months of collaborative depression care supported by a Web-based decision support system. However, given that medication was the primary mode of depression treatment within the trial, adherence to prescribed regimens was essential. Therefore, the aim of the current investigation was to isolate demographic, mental health, and physical health related factors that were associated with preintervention adherence patterns. Identification of these factors will be helpful for future efforts to improve provision of physical and mental health interventions for HIV-infected veterans.

Methods

Participants

Eligible participants were HIV-seropositive males and females (aged 18 and over), who were being treated for HIV at one of three Veterans Administration Medical Centers (VAMC) HIV clinics and were identified as having clinically significant symptoms of depression during the screening session (i.e., Patient Health Questionnaire [PHQ-9][21] depression score ≥ 10). Exclusion criteria were: no access to a telephone, current acute suicidal ideation, significant cognitive impairment as indicated by a score greater than 10 on the Blessed Orientation-Memory-Concentration (BOMC),[22] self-report history of bipolar disorder or manic depression, and medical record diagnosis of schizophrenia. After completing the informed consent process, participants completed the baseline assessment and were randomly assigned to intervention or usual care. Two hundred seventy-six individuals were randomized at the baseline research session and 225 are included in these analyses as they had a baseline prescription for either antiretroviral treatment ($n = 192$), antidepressant treatment ($n = 146$), or both ($n = 113$). All participants signed informed consents approved by their VAMC's Institutional Review Board (IRB) prior to the initiation of any research procedures.

Measures

Demographics The baseline interview included questions about the veteran's demographics characteristics. Data utilized in these analyses included gender, age, race, marital/partner status, and level of formal education.

Patient Health Questionnaire—9 items (PHQ-9)[21] The PHQ-9 is a 9-item self-report measure that was used to screen for the presence of symptoms of depression. A PHQ-9 score of

greater than 10 has strong psychometric properties in primary care settings (e.g., 99+% sensitivity and a 91% specificity).

Symptom Checklist—20 items (SLC-20).[23] Depression symptom severity over the 2 weeks before the baseline interview was measured using the Hopkins Symptom Checklist SCL-20. The SCL-20 includes the 13-item depression scale plus 7 depression-related items from the Hopkins Symptom Checklist-90–Revised. The items are scored from 0 to 4 and averaged to provide a mean depression severity score from 0 to 4.

Mental health diagnoses. The Mini International Neuropsychiatric Interview (MINI)[24] is a brief structured interview for the major Axis I psychiatric disorders, shown to be valid and reliable when compared to the Structured Clinical Interview for DSM-III-R and the CIDI.[25] The MINI was used to assess for the presence of clinically significant depression (i.e., major depression) and also for comorbid mental health conditions (i.e., generalized anxiety disorder, panic disorder, posttraumatic stress disorder, alcohol use disorder, any alcohol use over the last year and number of drinks in week before interview).

Physical comorbidity. Chronic physical health conditions (other than HIV) were measured using the 21-item Physical Comorbidity scale from the Depression Outcomes Module.[26,27]

HIV-related symptomatology. HIV symptom severity was measured using the 20-item Symptoms Distress Module,[28] which summarizes the degree to which each symptom bothered the participant in the past 4 weeks on a scale from 0 = I do not have this symptom to 4 = this symptom bothers me a lot. We also created a count variable to address the number of self-reported symptoms.

Quality of life. Health status was measured using the physical and mental health component summary scores from the Medical Outcomes Study SF-12V.[29] Health-related quality of life was measured using the Quality of Well-Being self-administered scale (QWB-SA).[30,31] The QWB-SA score is derived from general population preference weights and ranges from death (0.0) to perfect health (1.0).

Self-reported medication adherence. Antidepressant and HIV medication adherence were measured separately using the AIDS Clinical Trial Group assessment,[32] which asks participants to report the number of pills per day they are supposed to take and the number of pills they skipped taking for each medication for each of the past 4 days. Percent adherence was calculated as follows. First, for each of the last 4 days, the number of pills prescribed minus number of pills taken divided by the total number prescribed for each medication was calculated. Then, add the percentage adherence for each of the last 4 days and divide by 4; this algorithm allowed us to take into account the number of pills and to better account for the "weight" of a missed pill (e.g., missing 1 pill of a 2-pill regimen likely has a bigger influence on efficacy than missing 1 pill of a 6-pill regimen). However, the distributions for both antidepressant and antiretroviral adherence were skewed as approximately 75% of the sample reported 100% adherence. The distributions were not amenable to transformation and as such, we decided to dichotomize the data in "less than 90%" and "90% or greater." This cut-point was chosen as our recent work demonstrated that 90% adherence was the most sensitive cut-point for antidepressant adherence.[33] Although 95% adherence is a long accepted benchmark among HIV treatment adherence research,[34] more recent research indicates that level of adherence varies by regimen and resistance profile.[35,36] We chose 90% for both types of medication for consistency in analyses.

Patient knowledge of regimen. As the adherence data were based on self-report, we did a chart review for medications prescribed to examine patient knowledge of regimen. Previous research has identified that poor knowledge of one's regimen can be associated with nonadherence to HIV medications.[37] As such, we included this in our examination of possible correlates of adherence to both types of medication; however, due to limitations in our data, we were only able to compare knowledge of names of medications. We measured knowledge in two ways. First, we measured if they knew the correct number of antidepressant and antiretroviral medications that they were prescribed. Second, we coded the names of the medications so that we could compare patient's knowledge of both the number and names of their prescribed antidepressant and antiretroviral medications to the ones reflected in their chart (chart was assumed to be gold standard). We calculated a percentage by dividing the number of correctly identified medications by the number that were prescribed to them as noted in their chart (e.g., chart review said they were prescribed Med A, Med B, and Med C; Veteran reported Med A and Med C but NOT Med B; their knowledge percentage would be 66.7%.)

Procedure

Veterans were screened for depressive symptoms by clinic staff during routine clinical care visits with their HIV provider. Veterans who fulfilled the screening criteria were referred to research to learn about the larger HI-TIDES trial. If interested, they completed the informed consent procedure at that visit and were called by research staff, on average, 7 days later, completed their baseline interview. The baseline interview contained questions about demographics, physical and mental health symptoms, treatment history, treatment preferences, and self-reported treatment adherence for both antiretrovirals and antidepressants. This interview was completed before the veteran was informed about which arm they were randomized to (intervention versus TAU). Following completion of the baseline interview, research staff conducted a chart review of the participant's electronic medical record at the VA to gather information about their current prescriptions and comorbid health conditions.

Data analytic strategy

Analyses were conducted using SAS 9.2[38] (SAS Inc., Cary, NC) and Stata 9.0 (StataCorp, College Station, TX).[39] There were three steps in the analyses to identify predictors, run the seemingly unrelated bivariate probit (SUR) model, and finally, to add adherence as an explanatory variable. Preliminary analyses to examine the relationship between participant characteristics and adherence (separate analyses for antidepressant and antiretroviral adherence) were conducted using appropriate tests based on the distribution of the item.

Independent variables that were significant at the bivariate level $p < 0.20$ with either antidepressant or antiretroviral adherence were included as variables in the final models. To examine correlates of adherence, we used an SUR model, where the probability of adherence to HIV medications and the probability of adherence to antidepressant medications are modeled jointly. The SUR model specifically accounts for the possibility that the unmeasured factors affecting adherence to HIV medications also affect adherence to antidepressant medications, which leads to correlated error terms across the two probit regression equations. These unmeasured factors include omitted variables impacting adherence to antidepressant and HIV medications, and the measurement error common to assessing adherence to antidepressant medication and adherence to HIV medications. The SUR model assumes this correlation follows a bivariate normal distribution (with covariance ρ) and calculates whether the error terms are significantly correlated and the direction of the correlation.[40,41] The significance of the correlation coefficient ρ is tested using a likelihood ratio test that compares the log likelihood of the model where ρ is restricted to 0 to the log likelihood of the model where ρ is unrestricted. If the correlation is not statistically significant (e.g., $\rho = 0$), it implies that unmeasured factors influencing adherence to HIV medications and the unmeasured factors influencing adherence to antidepressant medications are different, and therefore that the two decisions are likely made independent of one another. In this case, the most appropriate statistical analysis involves estimating the parameters of two separate probit regressions. On the other hand, if the correlation is positive and significant, it implies that unmeasured factors affect the adherence to antiretroviral and antidepressant medications in the same way. Conversely, if the correlation is negative and significant, it suggests that the unmeasured factors affect adherence to HIV medications and antidepressant medications in opposite directions. In either case (significantly positive or negative correlated error terms), a significant correlation suggests that the decision to take HIV and antidepressant medications as prescribed are interdependent and should be modeled jointly using the seemingly unrelated bivariate probit model. Once the proper model specification (two independent probit models or seemingly unrelated bivariate probit model) was determined, significant correlates of nonadherence were examined based on the significance of the parameter estimates of the independent variables. The direction and magnitude of the correlate effects was determined by calculating standard marginal effects.[42] Finally, two additional models were run to examine the impact of each type of adherence on the other. When adherence for one type of medication is included as an explanatory variable in the SUR equation predicting adherence for the other type of medication, it is referred to as the bivariate probit with endogenous dummy model[41] or the recursive model for dichotomous choice.[43] For the model predicting antidepressant adherence to be fully identified, the probit equation includes one exogenous independent variable (depression severity) that is not included as independent variables in the probit equation predicting antiretroviral adherence. Likewise, for the model predicting antiretroviral adherence, to be fully identified, the probit equation includes one exogenous independent variable (HIV severity) that is not included as independent variables in the probit equation predicting antidepressant adherence (Fig. 1).

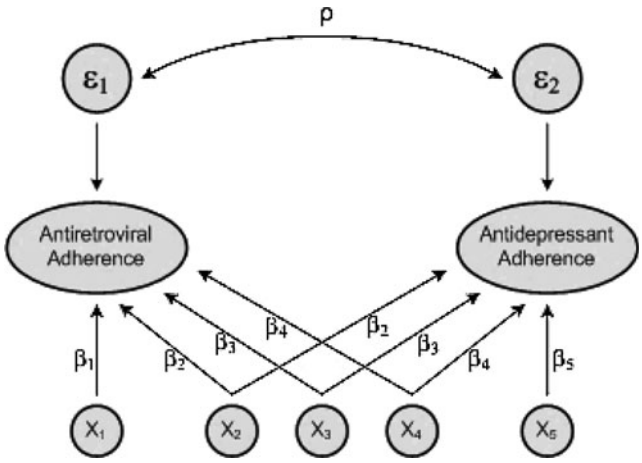


FIG. 1. Seeming unrelated regression (SUR) model. β, parameter estimates; ρ, correlation between the error terms; X, independent variables; ε, error terms.

Results

Sample characteristics

The sample consisted of 225 depressed HIV-infected veterans. Overall, the majority of the sample was male (97%), middle-aged (50 ± 10 years), and had a high school diploma (93%). Approximately 60% of the sample was African American, which is similar to the U.S. HIV population.[44] Over three fourths of the participants met criteria for major depressive disorder based on the MINI. Moreover, 75% of the sample also met criteria for at least 1 other comorbid mental health disorder. The participants reported a range of comorbid physical health conditions. The sample self-reported high rates of adherence to both their antidepressants and their antiretrovirals. Results indicated that 75.5% of those prescribed antiretrovirals ($n = 192$) reported 90% or greater adherence to their antiretroviral prescription while 76.7% of those prescribed antidepressants ($n = 146$) reported 90% or greater adherence to their antidepressant prescription (Table 1).

Among the subset of participants with both prescriptions ($n = 113$), adherence rates were as follows: both prescriptions less than 90%: $n = 11$ (9.73%); antiretroviral adherence less than 90% and antidepressant adherence 90% or greater: $n = 16$ (14.16%); antiretroviral adherence 90% or greater and antidepressant adherence less than 90%: $n = 10$ (8.85%); and both prescriptions 90% or greater : n = 76 (67.26%; see Table 2).

Table 3 provides descriptive characteristics and bivariate relationships between correlates and treatment adherence.

Bivariate relationships between correlates and treatment adherence

Separate analyses were conducted to examine bivariate relationships between demographic correlates, mental health diagnostic status, alcohol use measures, markers of physical health including HIV symptoms (both number and severity), measures of physical comorbidity, quality of life, medical knowledge, study-related factors (i.e., site), and treatment adherence (Table 3). Age ($p < 0.05$), race, education,

Table 1. Overview of Treatment Adherence and Medications Prescribed

	Antiretroviral medication (n = *192*) *Mean ± SD; range*	*Antidepressant medication* (n = *146*) *Mean ± SD; range*
Treatment adherence		
Antiretroviral adherence (%)	87.7 ± 25.7; range = 0–100	
Antidepressant adherence (%)		84.3 ± 32.3; range = 0–100
No. of medications prescribed		
Antiretrovirals—self-report	2.4 ± 1.1; range = 0–6	
Antiretrovirals—chart review	2.1 ± 1.3; range = 0–6	
Antidepressants—self-report		1.1 ± 0.4; range = 0–3
Antidepressants—chart review		0.9 ± 0.6; range = 0–3

SD, standard deviation.

depression severity, self-reported number ($p < 0.05$) and severity of HIV symptoms ($p < 0.05$), physical comorbidity, and quality of well-being ($p < 0.05$) were associated with antiretroviral adherence at $p < 0.20$ in bivariate analyses. Given the high correlation between the "number of HIV symptoms" variable and the "severity of HIV symptoms" variable, only the severity measure was used in subsequent analyses. In contrast, antidepressant adherence was associated with the following correlates at $p < 0.20$: gender, age ($p < 0.05$), marital status, generalized anxiety disorder diagnosis ($p < 0.05$), comorbid mental health disorder ($p < 0.05$), and any alcohol use in the past year. Given that the 'comorbid mental health condition' variable was highly correlated with "generalized anxiety disorder" variable, only the "generalized anxiety disorder" variable was used in subsequent analyses.

Final prediction model

A seemingly unrelated bivariate probit model was estimated and the positive correlation between the two errors terms was significant ($\rho = 0.64$, Wald test of ρ: $\chi^2(1) = 9.84$, $p = 0.002$). Therefore, correlates of adherence to antiretrovirals and antidepressants were estimated jointly. Education, age, and HIV symptom severity were significant correlates of antiretroviral medication adherence while gender and generalized anxiety disorder were significant correlates of adherence to antidepressant medications (see Table 4 for full results). The results from the bivariate probit with endogenous dummy models (available from first author upon request) indicate that antiretroviral adherence does not predict antidepressant adherence ($\beta = 1.62$, $p = 0.17$) but antidepressant adherence does predict antiretroviral adherence ($\beta = 2.30$, $p = 0.036$).

Table 2. Adherence Rates Among Those Prescribed Both Antidepressant and Antiretroviral Medications

	Antiretroviral medication	
Antidepressant medication	< 90% adherence	≥ 90% adherence
< 90% adherence	9.73%	8.85
≥ 90% adherence	14.16%	67.26%

Discussion

The present study attempted to isolate demographic, mental health, and/or physical health factors that were associated with preintervention adherence patterns among depressed HIV-infected patients. While antiretroviral adherence did not predict antidepressant adherence, antidepressant adherence predicted antiretroviral adherence. Antidepressant adherence predicting antiretroviral adherence is consistent with previous research.[5,16–18] Certain demographic factors (i.e., age and level of education), physical health (i.e., HIV severity), and mental health (i.e., generalized anxiety disorder diagnosis) were associated with adherence but there were relatively few significant correlates indicating the presence of omitted variables. We did note that older adults and those with comorbid generalized anxiety disorder were more adherent, suggesting that there may be additional attention to adherence with increasing age and also among those with a tendency to worry. The finding that a higher level of education was predictive of poorer adherence in the SUR equation predicting antiretroviral adherence was surprising. In a recent review of results from developed and developing nations, education was not identified as a predictor of antiretroviral adherence.[7] However, in our sample there was a relatively small number of subjects who did not have a high school diploma, therefore, this finding should not be over-interpreted. The negative relationship between adherence and HIV symptom severity suggests to us that those who are feeling better may be motivated to maintain this status via adherence to their regimens. Overall, our findings were modest, suggesting the variables we measured were of limited utility in predicting nonadherence behaviors.

The clinical implications of our findings may be helpful for clinical providers. Results suggested that antidepressant adherence affected adherence to antiretrovirals but not vice versa. Given that the subjects enrolled in this investigation are active patients in HIV clinics and were suffering from depression, this suggests that clinical providers' efforts to improve antidepressant adherence may also improve antiretroviral adherence, which is especially important given the virologic impact of inconsistent antiretroviral adherence. Our findings mirrors several other investigations that have also provided evidence of improved adherence to antiretrovirals among depressed individuals taking antidepressants.[5,16–18] However, we also simultaneously found that antiretroviral adherence did not impact antidepressant adherence. Even so, more research is needed to examine potential moderators

Table 3. Descriptive Characteristics and Bivariate Relationships Between Correlates and Treatment Adherence

Correlate variable	*Full sample* (n = *225*)	*Antiretroviral adherence* (n = *192*)		*Antidepressant adherence* (n = *146*)	
	Descriptive results	*Descriptive results*	*Bivariate relationship with correlate*	*Descriptive results*	*Bivariate relationship with correlate*
Demographics					
Site	Site 1: 143 (63.6%) Site 2: 56 (24.9%) Site 3: 26 (11.5%)	Site 1: 125 (65.1%) Site 2: 44 (22.9%) Site 3: 23 (12.0%)	$\chi^2(2) = 0.94$, $p = 0.63$	Site 1: 89 (61.0%) Site 2: 42 (28.8%) Site 3: 15 (10.2%)	$\chi^2(2) = 3.08$, $p = 0.21$
Gender	218 males (96.9%)	189 males (98.4%)	FET, $p = 0.57$	140 males (95.9%)	FET, $p = 0.14$
Age	49.9 ± 9.6 (range, 25–76)	50.2 ± 9.5 (range, 25–76)	$t(190) = -3.01$, $p < 0.01$	50.5 ± 9.8 (range, 26–76)	$t(144) = -3.09$, $p < 0.01$
African American	136 African Americans (60.4%)	113 African Americans (58.9%)	$\chi^2(1) = 1.70$, $p = 0.19$	81 African Americans (55.5%)	$\chi^2(1) = 0.63$, $p = 0.43$
Marital	179 single/not married (79.6%)	153 single/not married (79.7%)	$\chi^2(1) = 1.05$, $p = 0.31$	113 single/not married (77.4 %)	$\chi^2(1) = 2.98$, $p = 0.08$
Education level	208 high school graduate or more: (92.4%)	177 high school graduate or more: (92.2%)	FET, $p = 0.195$	135 high school graduate or more: (92.5%)	FET, $p = 1.00$
Mental health diagnosis					
Major depressive disorder	171 meet criteria: (76.0%)	145 meet criteria: (75.5%)	$\chi^2(1) = 0.04$, $p = 0.84$	120 meet criteria: (82.2%)	$\chi^2(1) = 0.0008$, $p = 0.98$
Generalized anxiety disorder	136 meet criteria: (60.4%)	117 meet criteria: (60.9%)	$\chi^2(1) = 0.32$, $p = 0.57$	95 meet criteria: (65.1%)	$\chi^2(1) = 6.32$, $p < 0.01$
Panic disorder	24 meet criteria: (10.7%)	20 meet criteria: (10.4%)	FET, $p = 0.27$	21 meet criteria: (14.4%)	FET, $p = 0.78$
Posttraumatic stress disorder	65 meet criteria: (28.9%)	49 meet criteria: (25.5%)	$\chi^2(1) = 0.15$, $p = 0.70$	53 meet criteria: (36.3%)	$\chi^2(1) = 0.91$, $p = 0.34$
Alcohol use disorder	38 meet criteria: (16.9%)	34 meet criteria: (17.7%)	$\chi^2(1) = 0.02$, $p = 0.89$	21 meet criteria: (14.4%)	FET, $p = 0.78$
Any alcohol use in last year?	166 drinks alcohol: (73.8%)	140 drinks alcohol: (72.9%)	$\chi^2(1) = 0.01$, $p = 0.92$	105 drinks alcohol: (71.9%)	$\chi^2(1) = 2.39$, $p = 0.12$
Number of drinks in week before baseline	3.6 ± 7.8; range, 0–53	3.6 ± 7.6; range, 0–53	Wilcoxon, $p = 0.92$	3.4 ± 7.6; range, 0–50	Wilcoxon, $p = 0.32$
Comorbid mental health condition?	168 meet criteria: (74.7%)	141 meet criteria: (73.4%)	$\chi^2(1) = 0.04$, $p = 0.84$	119 meet criteria: (81.5%)	$\chi^2(1) = 8.30$, $p < 0.01$
Depression					
SCL-20	1.8 ± 0.6; range, 0.4–3.6	1.8 ± 0.7; range, 0.4–3.6	$t(190) = 1.52$, $p = 0.13$	1.9 ± 0.6; range, 0.6–3.6	$t(144) = -0.48$, $p = 0.63$
Physical health					
Self-reported no. of HIV symptoms	11.1 ± 3.7; range, 3–19	10.8 ± 3.7; range, 3–19	Wilcoxon, $p < 0.001$	11.4 ± 3.6; range, 3–19	Wilcoxon, $p = 0.85$
Severity of self-reported HIV symptoms	35.7 ± 14.4; range, 6–76	34.8 ± 14.2; range, 6–76	$t(190) = 2.74$, $p < 0.05$	37.8 ± 14.0; range, 7–76	$t(144) = 0.13$, $p = 0.90$
Physical comorbidity	3.4 ± 2.3; range, 0–11	3.4 ± 2.4; range, 0–11	Wilcoxon, $p = 0.06$	3.5 ± 2.3; range, 0–11	Wilcoxon, $p = 0.63$
Quality of life					
Physical health component score of SF-12	40.9 ± 12.0; range, 16.1–64.9	41.5 ± 12.2; range, 16.1–64.9	$t(182) = -1.00$, $p = 0.32$	40.4 ± 12.1; range, 19.1–64.8	$t(139) = -0.49$, $p = 0.62$
Mental health component score of SF-12	35.0 ± 10.9; range, 10.8–61.2	35.1 ± 11.2; range, 10.8–61.2	$t(182) = 0.03$, $p = 0.98$	32.6 ± 9.9; range, 10.8–57.6	$t(139) = 0.31$, $p = 0.75$
Quality of well-being	0.5 ± 0.1; range, 0.2–0.7	0.5 ± 0.1; range, 0.2–0.7	$t(190) = -2.17$, $p = 0.03$	0.5 ± 0.1; range, 0.2–0.7	$t(144) = -0.70$, $p = 0.48$
Medication knowledge					
Knowledge variable		76.7 ± 36.2; range, 0–100	Wilcoxon, $p = 0.99$	70.6 ± 41.4; range, 0–100	Wilcoxon, $p = 0.21$

Note: FET = Fisher's Exact Test; χ^2 analysis was used for categorical data, with Fisher's exact test used when small cell counts rendered the χ^2 inappropriate. *t* tests were used for continuous data, with Wilcoxon used for non-normal distributions. Analyses with SF-12 are on subset due to missing data from SF-12 data collection.

Table 4. SUR with Significant Correlates ($p < 0.20$) and Treatment Adherence ($n = 113$)

	Antiretroviral adherence 90% cutoff		*Antidepressant adherence 90% cutoff*	
Correlate variable	*β (SE)*	*(95% CI)*	*β (SE)*	*(95% CI)*
Age	0.05 (0.02)[a]	0.01, 0.08	0.03 (0.02)	−0.01, 0.06
Gender	0.18 (0.80)	−1.39, 1.75	−5.60 (1.60)[a]	−8.74, −2.48
Race	0.20 (0.33)	−0.46, 0.85	−0.19 (0.30)	−0.79, 0.40
Education	−6.16 (1.55)[a]	−9.20, −3.12	−0.42 (0.63)	−1.66, 0.81
Marital	0.28 (0.36)	−0.42, 0.98	−0.03 (0.36)	−0.73, 0.67
Generalized anxiety disorder	0.27 (0.35)	−0.41, 0.94	0.67 (0.29)[a]	0.09, 1.25
Drink or abstain in last year?	0.17 (0.33)	−0.47, 0.82	−0.25 (0.32)	−0.87, 0.38
SCL-20	0.13 (0.28)	−0.42, 0.68	0.33 (0.29)	−0.24, 0.89
Severity of SR HIV symptoms	−0.03 (0.02)[a]	−0.07, −0.002	−0.01 (0.02)	−0.04, 0.02
Physical comorbidity	−0.09 (0.08)	−0.23, 0.06	0.01 (0.07)	−0.12, 0.15
Quality of well-being	−0.66 (1.31)	−3.23, 1.91	0.68 (1.43)	−2.12, 3.47

[a] $p < 0.05$; The correlation between the two errors terms of the seemingly unrelated bivariate probit model was significant ($\rho = 0.64$, Wald test of ρ: $\chi^2(1) = 9.84$, $p < 0.01$).

SUR, seemingly unrelated regression; CI, confidence interval.

of the relationship between antidepressant and antiretroviral adherence.

The present investigation has a number of strengths as well as noteworthy limitations. First, a major strength of the present investigation is the use of real-world patients, i.e., depressed HIV-infected veterans who were patients seeking care in three VA HIV clinics. The eligibility criteria for this investigation were minimal so as to facilitate generalizability of these finding and the overall trial to HIV-infected veterans seeking care in the VA more generally. Moreover, the racial makeup of our sample (60% African American) is similar to the U.S. HIV population,[44] which further improves generalizability. Second, we were able to examine correlates for both antiretroviral and antidepressant adherence simultaneously using the SUR modeling techniques. Limitations of the present investigation include reliance on self-reported treatment adherence (which involves considerable measurement error), limited variables available due to limitations in data collection, restriction of sample to veterans, use of cross-sectional data, few women in the sample, and modest sample size in the SUR model. Although there are benefits and drawbacks to self-reported HIV treatment adherence, it can be used to effectively estimate adherence.[45] Even so, future trials may want to build on our results using more objective methods of measurement such as MEMS caps, pill count, or pharmacy record review and utilize a window of measurement that is longer than 4 days.[46] Additionally, we believe that limitations in the variable "patient's knowledge of regimen" may be responsible for the nonassociation between the knowledge and adherence variables. Furthermore, though the Veterans Affairs network of health care (VAMCs) is the largest provider of HIV care in the United States, the veteran population does not mirror the overall HIV-infected population in the United States. Indeed, veterans in care for HIV infection are more likely to be male and older than the average U.S. HIV-infected adult. Although the percentage of female veterans receiving care at VAMCs is increasing, women are underrepresented in comparison to the percentage of HIV-infected women in the general U.S. population. We were limited in the interpretation of gender-related and possibly education-related findings given the limited variability of these factors in our sample. Finally, although we had 225 participants with either an antidepressant or antiretroviral prescription in our baseline sample, our sample size was much smaller ($N = 113$) when we restricted the sample to those with both prescriptions.

In conclusion, among our clinically depressed, treatment-seeking HIV-infected participants, demographic factors (i.e., age and level of education), physical health (i.e., HIV severity), and mental health (i.e., generalized anxiety disorder diagnosis) were correlates of self-reported adherence. We found limited support for additional mental and/or physical health factors as correlates of adherence to these regimens. Our findings also demonstrated that antidepressant adherence was associated with antiretroviral adherence in the SUR model but not vice versa. Future research is needed to identify and test additional factors and interventions that jointly impact antiretroviral and antidepressant adherence among treatment-seeking HIV-infected patients.

Author Disclosure Statement

No competing financial interests exist.

References

1. DiMatteo MR. Variations in patients' adherence to medical recommendations: A quantitative review of 50 years of research. Med Care 2004;42:200–209.
2. Ingersoll KS, Cohen J. The impact of medication regimen factors on adherence to chronic treatment: A review of literature. J Behav Med 2008;31:213–224.
3. Kacanek D, Jacobson DL, Spiegelman D, Wanke C, Isaac R, Wilson IB. Incident depression symptoms are associated with poorer HAART adherence: A longitudinal analysis from the Nutrition for Healthy Living study. J Acquir Immune Defic Syndr 2010;53:266–72.
4. Bulloch AGM, Patten SB. Non-adherence with psychotropic medications in the general population. Soc Psychiatry Psychiatr Epidemiol 2010;45:47–56.
5. Horzberg MA, Silverberg MJ, Hurley LB, et al. Effects of depression and selective serotonin reuptake inhibitor use on adherence to Highly Active Antiretroviral Therapy and on clinical outcomes in HIV-infected patient. J Acquir Immune Defic Syndr 2008;47:384–390.
6. Shacham E, Nurutdinova D, Onen N, Stamm K, Overton E. The interplay of sociodemographic factors on virologic

suppression among a U.S. outpatient HIV clinic population. AIDS Patient Care STDs 2010;24:229–235.
7. Mills EJ, Nachega JB, Bangsberg DR, et al. Adherence to HAART: A systematic review of developed and developing nation patient reported barriers and facilitators. PLoS Med 2006;3:e438.
8. Brown C, Battista DR, Bruehlman R, Sereika SS, Thase ME, Dunbar-Jacob J. Beliefs about antidepressant medications in primary care patients: Relationship to self-reported adherence. Med Care 2005;43:1203–1207.
9. Burra TA, Chen E, Mclntyre RS, Grace SL, Blackmore ER, Stewart DE. Predictors of self-reported antidepressant adherence. Behav Med 2007;32:127–134.
10. Chakraborty K, Avasthi A, Kumar S, Grover S. Attitudes and beliefs of patients of first episode depression towards antidepressants and their adherence to treatment. Soc Psychiatry Psychiatr Epidemiol 2009;44:482–488.
11. Raue P, Schulberg H, Heo M, Klimstra S, Bruce M. Patients' depression treatment preferences and initiation, adherence, and outcome: A randomized primary care study. Psychiatr Serv 2009;60:337–343.
12. Sheehan DV, Keene MS, Eaddy M, Krulewicz S, Kraus JE, Carpenter DJ. Differences in medication adherence and healthcare resource utilization patterns: Older versus newer antidepressant agents in patients with depression and/or anxiety disorders. CNS Drugs 2008;22:963–973.
13. Kaiser Permanente Care Management Institute. Depression Clinical Practice Guidelines. Oakland, CA: 0Kaiser Permanente Care Management Institute, 2006.
14. Sawada N, Uchida H, Suzuki T, et al. Persistence and compliance to antidepressant treatment in patients with depression: A chart review. BMC Psychiatry 2009;9:38.
15. Furtek K, Martin J, Burgi A, Crum-Cianfl N, Pope B, Hale B. Antiretroviral adherence and virologic response rates in HIV-infected patients with mood disorders [Abstract 994]. Poster presented at the 44th annual meeting of IDSA. Toronto: October 12–15, 2006.
16. Yun LWH, Marvai, M, Kobayashi JS, Barton PL, Davidson AJ. Antidepressant treatment improves adherence to antiretroviral therapy among depressed HIV-infected patients. J Acquir Immune Defic Syndr 2005;38:432–438.
17. Walkup J, Wei W, Sambamoorthi U, Crystal S. Antidepressant treatment and adherence to combination antiretroviral therapy among patients with AIDS and diagnosed depression. Psychiatr Q 2008;79:43–53.
18. Kumar V, Encinosa W. Effects of antidepressant treatment on antiretroviral regimen adherence among depressed HIV-infected patients. Psychiatr Q 2009;80:131–141.
19. Pyne JM, Fortney JC, Curran GM, et al. Effectiveness of collaborative care for depression in HIV clinics. Arch Intern Med 2011;171:23–31.
20. Fortney JC, Pyne JM, Edlund MJ, et al. A randomized trial of telemedicine-based collaborative care for depression. J Gen Intern Med 2007;22:1086–1093.
21. Kroenke K, & Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. Psychiatr Annu 2002;32: 509–515.
22. Katzman R, Brown T, Fuld P, Peck A, Schechter T, Schimmel H. Validation of a short orientation-memory-concentration test of cognitive impairment. Am J Psychiatry 1983;140:734–739.
23. Derogatis L. SCL-90-R. Administration, Scoring, and Procedures Manual, 3rd ed. Minneapolis, MN: National Computer Systems, Inc., 1994.
24. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. Eur Psychiatry 1997;12:232–241.
25. Lecrubier Y, Sheehan DV, Weiller E, et al. The Mini International Neuropsychiatric Interview (MINI): A short diagnostic structured interview: Reliability and validity according to the CIDI. Eur Psychiatry 1997;12:224–231.
26. Rost K, Smith G, Burnam M. Measuring the outcomes of care for mental health problems. Med Care 1992;30:MS266–MS273.
27. Smith GR, Jr., Burnam A, Burns BJ, Cleary P, Rost KM. Depression Outcomes Module (DOM). In: Handbook of Psychiatric Measures. Washington, D.C.: American Psychiatric Association, 2000:213–215.
28. Justice AC, Holmes W, Gifford AL, et al. Development and validation of a self-completed HIV symptom index. J Clin Epidemiol 2001;54(Suppl 1):S77–90.
29. Jones D, Kazis L, Lee A, et al. Health status assessments using the Veterans SF-12 and SF-36: Methods for evaluating outcomes in the Veterans Health Administration. J Ambul Care Manage 2001;24:68–86.
30. Kaplan RM, Anderson JP. The general health policy model: An integrated approach. In: Spiker B, ed. Quality of Life and Pharmacoeconomics in Clinical Trials. Second ed. Philadelphia: Lippincott-Raven Publishers, 1996:309–321.
31. Kaplan RM, Bush JW, Berry CC. Health status: Types of validity and the index of well-being. Health Serv Res 1976; 11:478–507.
32. Chesney MA, Ickovics JR, Chambers DB, et al. Self-reported adherence to antiretroviral medications among participants in HIV clinical trials: The AACTG adherence instruments. Patient Care Committee & Adherence Working Group of the Outcomes Committee of the Adult AIDS Clinical Trials Group (AACTG). AIDS Care 2000;12:255–266.
33. Fortney JC, Pyne JM, Edlund MJ, Mittal D. Relationship between antidepressant medication possession and treatment response. Gen Hosp Psychiatry 2010;32:377–379.
34. Paterson DL, Swindells S, Mohr J, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med 2000;133:21–30.
35. Bangsberg DR. Less than 95% adherence to nonnucleoside reverse-transcriptase inhibitor therapy can lead to viral suppression. Clin Infect Dis 2006;43:939–941.
36. Shuter J. Forgiveness of non-adherence to HIV-1 antiretroviral therapy. J Antimicrob Chemother 2008;61:769–773.
37. Miller LG, Liu H, Hays RD, et al. Knowledge of antiretroviral regimen dosing and adherence: A longitudinal study. Clin Infect Dis 2003;36:514–518.
38. SAS Institute Inc. SAS software, Version 9.2. Cary, NC, 2008.
39. StataCorp. Stata Statistical Software: Release 9. College Station, TX: StataCorp LP, 2005.
40. Costa-Font J, Gil J. Obesity and the incidence of chronic diseases in Spain: A seemingly unrelated probit approach. Econ Hum Biol 2005;3:188–214.
41. Fabbri D, Monfardini C, Radice R. Testing exogeneity in the bivariate probit model: Monte Carlo evidence and an application to health economics. Dipartimento di Scienze Economiche, Università di Bologna, Working Paper No.514; 2004.
42. Park HM. Regression Models for Binary Dependent Variables Using Stata, SAS, R, LIMDEP, and SPSS. Working Paper. The University Information TechnologyServices (UITS)

Center for Statistical and Mathematical Computing, Indiana University, 2009.
43. Maddala GS. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press, 1983.
44. VA Clinical Case Registry http://vaww.hiv.va.gov/ (Last accessed February 14, 2012).
45. Simoni J, Kurth A, Pearson C, Pantalone D, Merrill J, Frick P. Self-report measures of antiretroviral therapy adherence: A review with recommendations for HIV research and clinical management. AIDS & Behav 2006;10:227–245.
46. Lu M, Safren SA, Skolnik PR, et al. Optimal recall period and response task for self-reported HIV medication adherence. AIDS Behav 2008;12:86–94.

Address correspondence to:
Kathryn A. Bottonari, Ph.D.
Ralph H. Johnson VA Medical Center
109 Bee Street
Charleston, SC 29401

E-mail: Kathryn.Bottonari@va.gov

ORIGINAL RESEARCH

American Heart Association | American Stroke Association

Prevalence, Distribution, and Risk Factor Correlates of High Thoracic Periaortic Fat in the Framingham Heart Study

Kathryn A. Britton, MD, MPH; Alison Pedley, PhD; Joseph M. Massaro, PhD; Erin M. Corsini, BS; Joanne M. Murabito, MD, ScM; Udo Hoffmann, MD, MPH; Caroline S. Fox, MD, MPH

***Background*—**Thoracic periaortic adipose tissue (TAT) is associated with atherosclerosis and cardiovascular disease (CVD) risk factors and may play a role in obesity-mediated vascular disease. We sought to determine the prevalence, distribution, and risk factor correlates of high TAT.

***Methods and Results*—**Participants from the Framingham Heart Study (n=3246, 48% women, mean age 51.1 years) underwent multidetector computed tomography; high TAT and visceral adipose tissue (VAT) were defined on the basis of sex-specific 90th percentiles in a healthy referent sample. The prevalence of high TAT was 38.1% in women and 35.7% in men. Among individuals without high VAT, 10.1% had high TAT. After adjustment for age and VAT, both women and men with high TAT in the absence of high VAT were older and had a higher prevalence of CVD ($P<0.0001$) compared with those without high TAT. In addition, men in this group were more likely to be smokers ($P=0.02$), whereas women were more likely to have low high-density lipoprotein cholesterol ($P=0.005$).

***Conclusions*—**Individuals in our community-based sample with high TAT in the absence of high VAT were characterized by an adverse cardiometabolic profile. This adipose tissue phenotype may identify a subset of individuals with distinct metabolic characteristics. **(*J Am Heart Assoc.* 2012;0:e004200 doi: 10.1161/JAHA.112.004200)**

Key Words: body fat distribution • obesity • perivascular adipose tissue • risk factors • visceral adipose tissue

Obesity is associated with cardiovascular morbidity and mortality.[1] The mechanisms by which obesity might contribute to vascular disease remain incompletely understood. Body fat distribution may be a cardiovascular risk factor even after accounting for generalized adiposity.[2] One component of abnormal body fat deposition involves the deposition of adipose tissue, so-called ectopic fat, around organs and the vasculature.[3] Perivascular fat is one such ectopic fat depot that has been postulated to have a local pathogenic effect on blood vessels.[4–8] Thoracic periaortic fat is a subtype of perivascular fat that can be quantified using multidetector computed tomography (MDCT).[9]

Thoracic periaortic fat (TAT) may be a novel risk marker for cardiovascular disease.[10,11] We have previously shown that TAT is associated with certain metabolic risk factors after adjustment for body mass index (BMI) as well as abdominal aortic and coronary calcium after adjustment for either BMI or visceral adipose tissue (VAT).[11] However, the prevalence and age distribution of TAT in a community-based sample has not been described. In addition, given the known correlation between TAT and VAT[11] and the known association of VAT with cardiometabolic risk,[12] we sought to examine the relative association of high TAT versus high VAT with cardiometabolic risk by examining different patterns of thoracic periaortic and visceral fat deposition in a cohort of middle-aged individuals enrolled in the Framingham Heart Study.

Methods

In 1971, children of those in the original Framingham Heart Study cohort and their spouses were enrolled in the

From the Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, MA (K.A.B.); Harvard Medical School, Boston, MA (K.A.B., U.H., C.S.F.); National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA (K.A.B., A.P., J.M. Murabito, C.S.F.); Department of Mathematics and Statistics, Boston University, Boston, MA (J.M. Massaro); Department of Radiology, Massachusetts General Hospital, Boston, MA (E.M.C., U.H.); Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA (J.M. Murabito); Division of Intramural Research and the Center for Population Studies, National Heart, Lung, and Blood Institute, Framingham, MA (C.S.F.); Division of Endocrinology, Metabolism and Hypertension, Brigham and Women's Hospital and Harvard Medical School, Boston, MA (C.S.F.).

Correspondence to: Caroline Fox, MD, MPH, NHLBI's Framingham Heart Study, 75 Mt. Wayte Avenue, Suite 2, Framingham, MA 01702-5827. E-mail: foxca@nhlbi.nih.gov

This article was handled independently by Viola Vaccarino, MD, PhD, as Guest Editor.

Received July 23, 2012; accepted October 17, 2012.

Offspring cohort. In 2002, individuals with at least 1 parent in the Offspring cohort were enrolled in the Third-Generation cohort. The study designs have been described previously.[13,14] The present study included participants from the Offspring and Third-Generation cohorts who participated in the MDCT substudy between 2002 and 2005 as previously described.[12] Of the 3529 participants in the MDCT substudy, 3246 had interpretable values for TAT. Of the individuals with interpretable TAT, 3228 had interpretable values for VAT. For individual regressions, any individuals with data missing the covariate of interest were excluded. The study protocol was approved by the institutional review boards of the Boston University Medical Center and Massachusetts General Hospital. All subjects provided written informed consent.

MDCT Scan Protocol and Adipose Tissue Measurements

Participants underwent radiographic assessment of their thorax and abdomen in the supine position using an 8-slice MDCT scanner (LightSpeed Ultra, General Electric, Milwaukee, WI) as previously described.[9] The thoracic scan was performed during an inspiratory breath hold with prospective ECG triggering (with the center of acquisition window at 70% of the R-R cycle to minimize cardiac motion). The average scan time was 18 seconds (tube voltage of 120 kVp, tube current of 320 mA (<220 lbs) or 400 mA (>220 lbs) with a gantry rotation time of 500 ms and a temporal resolution of 330 ms. Thoracic and abdominal MDCT images were reconstructed as 2.5- and 5-mm nonoverlapping slices, respectively.

TAT and VAT were assessed using a dedicated workstation (Aquarius 3D, TeraRecon, San Mateo, CA). Fat volumes were measured by a semiautomatic segmentation technique requiring manual definition of tissue borders. Fat within an area of interest was defined by pixels with characteristic Hounsfield units (HU; window width −195 to −45 HU; window center −120 HU). The area of interest for TAT was defined anteriorly by the area immediately surrounding the thoracic aorta (defined by a line drawn horizontally through the esophagus, which connected to the left costovertebral joint) and posteriorly by the right lateral border of the vertebral body and the anterior edge of the vertebral body.[9] This resulted in a 6.75-cm column of fat (27 slices) surrounding the thoracic aorta. VAT was quantified on abdominal scans as previously described.[12] Briefly, the reader defined the area of interest by tracing the abdominal muscular wall and separating the subcutaneous from the visceral abdominal fat depot. These areas of interest were summed over the 25 abdominal slices. Intrareader and interreader (assessed between 2 readers) reproducibility was excellent for both TAT and VAT, with intraclass and interclass correlations >0.98.[9,15] A visual representation of the methodology for measuring TAT is presented in Figure 1.

Covariate Assessment

Covariates were measured at the seventh Framingham Offspring examination (1998–2001) and the first Third-Generation examination (2002–2005). BMI was defined as weight (in kilograms) divided by the square of the height (in meters). Waist circumference (WC) was measured at the level of the umbilicus. Current smoking was defined as smoking ≥1 cigarette per day in the past year. Alcohol use was assessed by physician-administered questionnaires, and dichotomized on the basis of consumption of ≥14 drinks per week (in men) or ≥7 drinks per week (in women). A physical activity index (PAI) score was calculated by summing the reported numbers for each level of activity, weighted by their estimated metabolic expenditure, as described previously.[16] The PAI ranges from a minimum score of 24, indicating 24 hours of

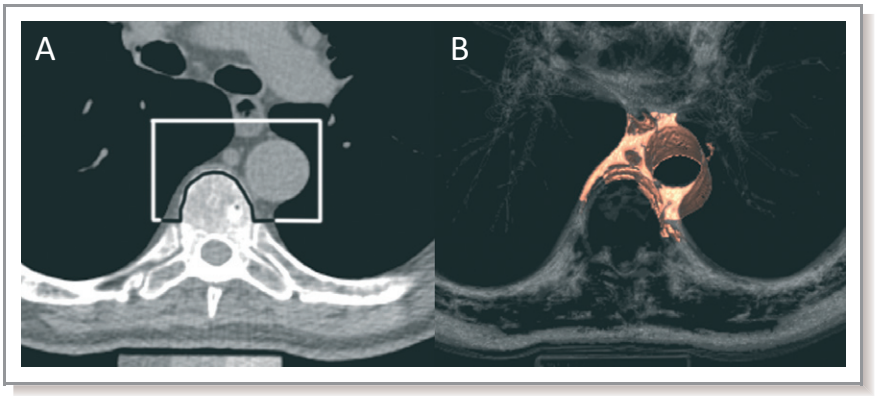


Figure 1. A, The region of interest drawn around the aorta using anatomic landmarks. Adipose tissue within this region of interest, defined as pixels with Hounsfield units between −195 and −45, is considered periaortic fat. B, 3D reconstruction. Reproduced from Fox et al (10) with permission from the publisher.[10]

sleeping, to a maximum score of 120, indicating 24 hours of heavy physical activity. Serum triglycerides, total and high-density lipoprotein cholesterol, and fasting plasma glucose were measured on fasting morning samples. Fasting plasma glucose ≥126 mg/dL or treatment with a hypoglycemic agent or insulin was used to define diabetes mellitus. Hypertension was defined as a systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg or treatment with an antihypertensive agent. Metabolic syndrome was defined from the modified Adult Treatment Panel criteria.[17] If periods had stopped for >1 year, women were considered postmenopausal. Cardiovascular disease included coronary heart disease, stroke, intermittent claudication, and congestive heart failure.

Statistical Analysis

To determine the prevalence of high TAT, a healthy referent sample was created by hierarchical exclusion of participants with the following covariates: BMI≥30 kg/m^2 (n=874), presence of hypertension (or use of antihypertensive medications; n=562), triglycerides ≥150 mg/dL or lipid-lowering medication (n=424), low high-density lipoprotein (HDL) cholesterol (<40 mg/dL in men and <50 mg/dL in women; n=193), impaired fasting glucose, diabetes or use of hypoglycemic medications (n=215); prevalent cardiovascular disease (CVD; n=9); current tobacco smoking (n=118); BMI<18.5 kg/m^2 (n=12), and missing covariates (n=24), resulting in 499 women and 316 men. We defined high TAT as a sex-specific fat volume ≥90th percentile from the healthy referent sample. Sex-specific TAT volumes <90th percentile from the healthy referent sample were classified as normal. The 90th percentile was chosen to ensure adequate sample size in the healthy referent sample to provide statistically robust estimates for the fat volume cutoff values. This method used to determine the volume for high TAT was the same as that used for other fat depots in the Framingham Heart Study.[18,19] High TAT, by this definition, corresponded with volumes ≥10.2 cm^3 in women and ≥19.0 cm^3 in men. Prevalence estimates of high TAT were then determined in the overall sample (including those individuals who had been excluded from the healthy referent sample) and stratified by sex and the following age categories: 35 to 44, 45 to 54, 55 to 64, 65 to 74, and 75 to 84 years.

To establish the risk factor profiles associated with different patterns of TAT and VAT, we stratified our sample into 4 mutually exclusive categories on the basis of TAT and VAT volumes: (A) normal TAT and normal VAT, (B) high TAT but normal VAT, (C) high VAT and normal TAT, and (D) high TAT and high VAT. High VAT has been previously defined as ≥1359 cm^3 in women and ≥2323 cm^3 in men. Given the known correlation between TAT and VAT,[10] we examined differences in risk factor levels between the discordant categories (groups B versus C) using sex-specific age-adjusted analyses of covariance (ANCOVAs) and logistic regressions. Given that high VAT is already recognized as associated with an adverse cardiometabolic risk factor profile, we also examined differences in risk factors among individuals with normal VAT that were discordant for high TAT (groups A versus B). For these analyses, we used sex-specific ANCOVAs and logistic regressions adjusted for age as well as the volume of VAT. We also adjusted for VAT given that individuals with high TAT and normal VAT tended to have higher absolute volumes of VAT compared with those with normal TAT and normal VAT. We did not adjust for height given the lack of documented association between height and TAT (r=0.05, P=0.25 in women; r=0.01, P=0.75 in men)[11] We tested for an age interaction across the 4 different categories of TAT/VAT. When age interactions were significant, we performed additional analyses stratified by median age. Finally, we performed a sensitivity analysis restricting our sample to the Third-Generation cohort in whom risk factor assessment and MDCT scans were performed during the same period.

SAS version 9.1 was used to perform all computations. Two-sided P values <0.05 were considered significant. Because of the exploratory nature of this study, no adjustments were made for multiple comparisons.

Results

Overall, 1558 women and 1688 men for a total sample of 3246 individuals were included in this analysis. For the overall sample, the mean values for TAT in women and men were 9.9 and 17.5 cm^3, respectively. The characteristics of the sample population are described in Table 1.

Distribution of Thoracic Periaortic Fat in the Community

Sex- and age-specific TAT volumes are represented in Table 2. TAT volumes increased with age into the eighth decade for women and into the seventh decade for men for all percentiles of fat.

Prevalence of High Thoracic Periaortic Fat

The overall and age-stratified prevalence of high TAT, defined as ≥90th percentile in the healthy referent sample, is reported in Table 3. The overall prevalence for high TAT was 38.1% in women and 35.7% in men and increased with age (P value for linear trend <0.0001). As expected, the prevalence of high TAT increased with increasing BMI and waist circumference categories (Table 4).

Table 1. Clinical Characteristics of the Overall Sample*

Characteristic	Overall (N=3246)	Women (n=1558)	Men (n=1688)
Age, y	51.1±10.4	52.3±9.9	49.9±10.7
BMI (kg/m^2)	27.7±5.2	27.0±5.8	28.4±4.5
Waist circumference (cm)	97.1±14.2	93.0±15.5	100.8±11.7
Low HDL cholesterol† (%)	29	26	33
Elevated triglycerides‡ (%)	36	27	44
Smoking status (%)			
Current	13	12	13
Former	39	43	37
Never	48	45	50
Alcohol use§ (%)	15	15	16
Physical activity index	37.6±7.2	36.7±5.8	38.3±8.2
Postmenopausal (%)	—	51	—
Hypertension (%)	29	27	32
Impaired fasting glucose‖ (%)	29	19	38
Diabetes (%)	7	6	7
Metabolic syndrome (%)	33	27	38
CVD (%)	6	4	8
Hypertension treatment (%)	19	19	20
Lipid treatment (%)	14	10	18
TAT (cm^3)	13.8±8.2	9.9±5.3	17.5±8.7
VAT (cm^3)	1818.7±1034.2	1356.7±825.3	2243.2±1025.0

Data are presented as mean±SD for continuous or % for categorical characteristics. BMI indicates body mass index; HDL, high-density lipoprotein; CVD, cardiovascular disease; TAT, thoracic periaortic fat; VAT, visceral adipose tissue.
*Sample sizes vary from row to row, as available data were used for a given characteristic.
†Defined as <40 mg/dL (men) and <50 mg/dL (women).
‡Defined as ≥150 mg/dL or lipid treatment.
§Defined as ≥14 drinks weekly for men and ≥7 drinks weekly for women.
‖Defined as fasting plasma glucose 100 to 125 mg/dL and not currently taking diabetes medication.

Cardiometabolic Risk Factor Profiles by Patterns of Thoracic Periaortic Fat and Visceral Adipose Tissue

Given the known high correlation between TAT and VAT (r=0.75, P<0.001),[11] we divided individuals into 4 categories on the basis of the presence or absence of high TAT and the presence or absence of high VAT to better evaluate the unique correlates of each fat depot. We then compared cardiometabolic risk factors between individuals with high TAT and normal VAT versus high VAT and normal TAT. In addition, among individuals with normal VAT, we compared cardiometabolic risk factors among individuals discordant for high TAT.

In the overall sample, 16.4% of women and 18.8% of men were discordant for high TAT and high VAT. Comparing discordant groups, women with high TAT and normal VAT compared with high VAT and normal TAT were older but tended to have lower measures of clinical adiposity and better cardiometabolic risk factor profiles after adjusting for age (Table 5). In contrast, men with high TAT and normal VAT compared with high VAT and normal TAT were older and had a lower waist circumference, but otherwise the cardiometabolic profiles were not significantly different (Table 5).

Among individuals with normal VAT, high TAT compared with normal TAT was associated with a more adverse cardiometabolic profile. Specifically, these individuals were older and had a higher prevalence of the majority of cardiometabolic risk factors, including a higher prevalence of metabolic syndrome (Table 5). Findings were similar when we also adjusted for BMI (data not shown). However, individuals with high TAT and normal VAT also had a higher absolute volume of VAT compared with individuals with normal TAT and normal VAT. Given this difference in VAT volumes, our findings presented in Figure 2 reflect additional adjustment of these models for the absolute volume of VAT. In these models, the presence of high TAT was associated with prevalent CVD in both sexes (age- and VAT-adjusted P=0.01 in women, Figure 2A; and P=0.004 in men, Figure 2B). In addition, among women with normal VAT, high TAT was associated with significantly lower HDL levels (P=0.005) (Figure 2A). In men with high TAT and normal VAT, there was a higher prevalence of smoking (age-VAT-adjusted P=0.02) and higher BMI (P=0.004) compared with individuals with normal VAT and normal TAT. After adjustment for the volume of VAT, the difference in the prevalence of metabolic syndrome was no longer different between those with high TAT versus normal TAT. When we reexamined the association of high TAT versus normal TAT with CVD prevalence after adjustment for low HDL and current smoking (in addition to adjustment for age and volume of VAT), there remained a significant association in both women (P=0.02) and men (P=0.01).

There were multiple significant age interactions across the 4 categories of TAT/VAT categories. Thus, we performed analysis stratified by median age. The results were generally similar to the overall findings (data not shown). Finally, in our sensitivity analysis limited to the Third Generation cohort, who underwent risk factor and MDCT assessment during the same time period, findings were overall similar (Table 6).

Table 2. Thoracic Periaortic Fat Percentiles Within Age Groups Among Women (n=1543) and Men (n=1634)*

	Fat Volumes, cm^3							
	N	5th	10th	25th	50th	75th	90th	>95th
Women (age), y								
35 to 44	391	3.2	3.7	4.8	6.3	8.8	11.4	13.8
45 to 54	591	3.9	4.4	5.8	8.1	11.3	15.4	17.7
55 to 64	341	5.2	6.0	7.9	10.8	14.1	17.7	20.6
65 to 74	185	6.4	7.1	9.7	13.1	17.7	22.6	25.1
75 to 84	35	7	7.6	11.0	14.5	19.3	26.5	28.2
Men (age), y								
35 to 44	559	5.9	6.7	9.5	12.5	16.9	22.0	26.8
45 to 54	582	7.2	8.9	11.7	15.2	20.1	25.2	29.5
55 to 64	280	9.8	11.7	15.1	20.2	26.3	31.8	38.7
65 to 74	182	11.1	12.6	17.5	23.6	31.7	39.6	43.9
75 to 84	31	7.9	13.2	16.2	23.1	34.5	42.7	43.9

*Fifteen women and 54 men from the total sample were either <35 years of age or missing information on age and therefore were excluded from this analysis.

Table 3. Sex-Specific Prevalence (Standard Error) of Excess TAT* by Age Group

	Women (n=1543)†	Men (n=1634)†
Overall	38.1% (1.24)	35.7% (1.19)
35 to 44	15.3% (1.82)	17.9% (1.62)
45 to 54	32.7% (1.93)	29.4% (1.89)
55 to 64	52.8% (2.70)	57.5% (2.95)
65 to 74	69.2% (3.39)	72.0% (3.33)
75 to 84	77.1% (7.10)	67.7% (8.40)
P value for linear trend	<0.0001	<0.0001

TAT indicates thoracic periaortic fat.
*Excess TAT defined as ≥90 percentile of the sex-specific cut points (≥10.2 and ≥19.0 cm^3 in women and men, respectively) in a healthy referent sample.
†Fifteen women and 54 men from the total sample were either <35 years of age or missing information on age and therefore were excluded from this analysis.

Discussion

In our community-based sample, more than a third of individuals had high TAT. In the absence of high VAT, excess TAT identified a subset of individuals with a higher prevalence of adverse cardiometabolic characteristics compared with individuals without high TAT even after adjustment for total volume of VAT. This included a higher prevalence of CVD in both women and men, lower HDL levels in women, and a higher prevalence of current smoking in men. These findings provide a better understanding of the risk factor correlates of TAT, a specific subtype of perivascular fat. Further experimental studies are required to elucidate whether TAT is pathogenic.

Multiple basic science and small clinical studies have suggested a local effect of perivascular fat that changes with the development of obesity.[5,6,8] Previous work in the

Table 4. Sex-Specific Prevalence (Standard Error) of High TAT* by BMI and Waist Circumference Category

BMI Category	Women (n=1544)	Men (n=1674)	Waist Circumference Category	Women (n=1539)	Men (n=1671)
Normal weight (BMI<25 kg/m^2)	12.4% (1.26)	6.6% (1.30)	Normal waist circumference (≤88 cm in women, ≤102 cm in men)	9.8% (1.16)	16.2% (1.18)
Overweight (25 kg/m^2≤BMI<30 kg/m^2)	43.7% (2.28)	29.7% (1.59)	High waist circumference (>88 cm in women, >102 cm in men)	58.5% (1.66)	60.8% (1.85)
Obese (BMI≥30 kg/m^2)	75.4% (2.18)	65.0% (2.17)			

TAT indicates thoracic periaortic fat; BMI, body mass index.
*High TAT defined as ≥90 percentile sex-specific cut points (≥10.2 and ≥19.0 cm^3 in women and men, respectively) for TAT in a healthy referent sample.

Table 5. Age-Adjusted Sex-Specific Distribution of Risk Factors and Clinical Characteristics by VAT/TAT Categories* Among Women (n=1546) and Men (n=1682)†

Risk Factor	Normal VAT/ Normal TAT (A)	High VAT/ Normal TAT (B)	High TAT/ Normal VAT (C)	High VAT/ High TAT (D)	Age-Adjusted *P* Value Comparing Groups B and C
Women					
n	787	168	86	505	
Age (y)	48.7	51.6	57.3	57.2	<0.0001
BMI (kg/m^2)	23.4	29.2	26.2	32.1	<0.0001
Waist circumference (cm)	83.4	99.6	89.8	106.6	<0.0001
VAT (cm^3)	750.5	1693.6	1082	2236.4	<0.0001
TAT (cm^3)	6.5	8.6	12.1	15.1	<0.0001
Diabetes (%)	2.6	5.8	4.0	8.8	0.50
Impaired fasting glucose‡ (%)	8.6	24.2	12.8	32.7	0.03
Hypertension (%)	16.6	31.7	21.1	39.5	0.049
Elevated triglycerides§ (%)	12.8	34.9	28.3	44.1	0.27
Low HDL cholesterol‖ (%)	12.3	35.0	27.7	44.7	0.25
Metabolic syndrome (%)	5.9	38.0	17.4	55.6	0.0006
CVD (%)	1.5	4.0	7.5	6.2	0.20
Smoking status (%)					
Current	9.9	12.3	17.4	16.7	0.29
Former	42.2	37.0	41.9	45.4	0.44
Never	48.2	50.8	41.1	38.7	0.15
Men					
n	876	216	101	489	
Age (y)	46	49.4	55.4	56.0	<0.0001
BMI (kg/m^2)	25.9	29.6	29.0	32.3	0.17
Waist circumference (cm)	94.2	104.7	102.0	110.6	0.017
VAT (cm^3)	1525.4	2788.6	1994.5	3340.6	<0.0001
TAT (cm^3)	12.2	15.4	22.4	26.6	<0.0001
Diabetes (%)	3.7	9.0	7.0	11.4	0.53
Impaired fasting glucose‡ (%)	29.7	40.3	39.8	49.5	0.92
Hypertension, %	22.5	33.0	35.7	45.4	0.62
Elevated triglycerides§ (%)	31.1	57.3	56.2	57.7	0.85
Low HDL cholesterol‖ (%)	22.5	41.7	40.2	46.2	0.80
Metabolic syndrome (%)	17.0	51.8	45.7	66.9	0.32
CVD (%)	4.6	8.5	11.7	9.6	0.32
Smoking status (%)					
Current	11.3	14.0	18.6	16.3	0.31
Former	33.1	35.5	38.8	42.1	0.55
Never	55.5	50.9	43.0	41.4	0.19

VAT indicates visceral adipose tissue; TAT, thoracic periaortic fat; BMI, body mass index; HDL, high-density lipoprotein; CVD, cardiovascular disease.

*TAT and VAT categories are defined as high if ≥90th percentile sex-specific cut points in healthy referent sample. High TAT corresponded to volumes ≥10.2 and ≥19.0 cm^3 in women and men, respectively. High VAT has been previously defined as ≥1359 cm^3 in women and ≥2323 cm^3 in men.

†Sample sizes vary from row to row, as available data were used for a given characteristic.

‡Defined as fasting plasma glucose 100 to 125 mg/dL and not currently taking diabetes medication.

§Defined as ≥150 mg/dL or lipid treatment.

‖Defined as <40 mg/dL (men) and <50 mg/dL (women).

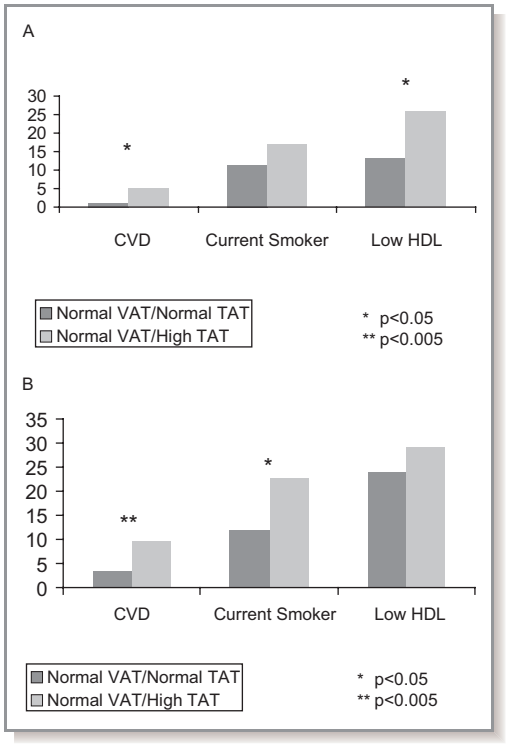


Figure 2. Differences in cardiometabolic features among (A) women (n=873) and (B) men (n=977) with normal VAT/normal TAT vs normal VAT/high TAT. Prevalence rates and *P* values are adjusted for age and volume of VAT. CVD indicates cardiovascular disease; HDL, high-density lipoprotein; TAT, thoracic periaortic adipose tissue; VAT, visceral adipose tissue.

Framingham Heart Study has demonstrated a novel and reliable method of quantifying thoracic periaortic fat by MDCT.[9] We now present a comprehensive age- and sex-specific description of TAT distribution in a community-based sample and report differences in cardiometabolic features among subcategories of TAT and VAT.

Our findings that more than one third of individuals in the community had high TAT is similar to our previous reports of fat depots, including VAT, pericardial, and intrathoracic fat.[18,19] Despite this similar prevalence among various fat depots, our findings suggest that a high volume of TAT is associated with adverse cardiometabolic features among the subset of individuals with normal VAT. Prior work has highlighted the existence of "metabolically obese normal weight individuals" who exhibit glucose intolerance and hyperinsulinemia despite a normal BMI.[20] Differences in visceral adiposity have been postulated to contribute to this phenotype.[21] Our findings suggest that high thoracic periaortic fat may also identify a "metabolically obese" group among individuals who do not meet criteria for excess VAT. Some of these differences, such as the higher prevalence of CVD, are present in both sexes. However, others are sex specific, with a prevalence of low HDL higher in women than men. This finding is consistent with prior literature suggesting stronger associations of ectopic fat and metabolic risk factors in women compared with men. However, the prevalence of current smoking was higher in men than in women with high TAT/normal VAT compared with normal TAT/normal VAT.

TAT is directly wrapped around the aorta, and this distinct anatomic location may explain the specific association between high TAT and CVD among individuals with normal VAT. TAT may serve as a marker of perivascular fat throughout the body including smaller blood vessels, and perivascular fat has been postulated to have adverse effects on the vasculature.[4,5] Supporting this, prior work in the Framingham Heart Study has demonstrated an association between TAT and both abdominal aortic calcium and coronary artery calcium among individuals without known cardiovascular disease. These associations persisted after adjustment for VAT and standard cardiovascular risk factors.[10] Alternatively, our findings that high TAT was associated with CVD among individuals with normal VAT may reflect that these individuals were more likely to have a higher prevalence of certain CVD risk factors, including low HDL in women and smoking in men.

Our findings of a difference in the prevalence of smoking among men with normal VAT but discordant for TAT deserves specific comment. Smokers tend to have a lower body weight but are known to have a higher risk of cardiovascular disease. Ectopic fat distribution, specifically higher volumes of VAT, is already known to differ between smokers and nonsmokers.[22] We have extended these findings among individuals with normal VAT by demonstrating a higher prevalence of male smokers among individuals with high versus normal TAT. These findings are noteworthy given the potential modulating role of nicotine on perivascular fat previously demonstrated in animal models.[23] Specifically, exposure of rats to nicotine prenatally and during lactation led to an increase in total adiposity as well as perivascular fat in the offspring compared with controls. Furthermore, the normal anticontractile effect of perivascular fat on blood vessels was no longer present in nicotine-exposed offspring, but was restored with the transfer of the medium surrounding normal fat. Thus, nicotine was associated with higher volumes as well as dysfunction of perivascular fat. These findings are consistent with prior work suggesting that excess perivascular adipose tissue disrupts the normal contribution of perivascular fat to vascular tone[24] and further suggest that nicotine may contribute to this process. Other studies also support a potential role of nicotine on adipose tissue.[25,26] For example, differences in adipose tissue lipoprotein lipase activity have been found between smokers and nonsmokers.[25] Thus, elucidation of the full expression profile of perivascular fat in response to nicotine may help to extend our observational findings.

Table 6. Age-Adjusted Sex-Specific Distribution of Risk Factors and Clinical Characteristics by VAT/TAT Categories* Among Women (n=848) and Men (n=1038) Limited to the Third Generation†

Risk Factor	Normal VAT/ Normal TAT (A)	High VAT/ Normal TAT (B)	High TAT/ Normal VAT (C)	High VAT/ High TAT (D)	Age-Adjusted *P* Value Comparing Groups B and C
Women					
n	518	54	66	210	
Age (y)	45.3	46.2	47	48.2	0.40
BMI (kg/m^2)	23.4	30.2	26.3	33.2	<.0001
Waist circumference (cm)	82.3	99.8	90.9	108.2	<.0001
VAT (cm^3)	688.8	1653.8	1037.9	2125	<.0001
TAT (cm^3)	5.7	7.5	10.2	13.1	<.0001
Diabetes (%)	1.5	5.6	1.4	6.1	0.21
Impaired fasting glucose‡ (%)	5.9	24.1	10.2	30.4	0.04
Hypertension (%)	10.0	25.9	15.2	31.1	0.13
Elevated triglycerides§ (%)	6.6	22.2	25.1	39.1	0.71
Low HDL cholesterol‖ (%)	14.5	38.9	24.7	47.3	0.10
Metabolic syndrome (%)	2.8	28.2	14.5	52.3	0.06
CVD (%)	0.2	1.9	0.0	3.0	0.95
Smoking status					
Current	11.7	16.7	15.8	19.9	0.90
Former	36.6	27.8	41.6	36.2	0.12
Never	51.9	55.6	42.5	44.4	0.16
Men					
N	585	85	130	238	
Age (y)	42.5	44.4	45.5	46.4	0.22
BMI (kg/m^2)	25.9	30	28.1	32.7	0.0001
Waist circumference (cm)	93.3	104.7	100.4	111.5	0.0007
VAT (cm^3)	1424.7	2744.4	1940.7	3094.9	<.0001
TAT (cm^3)	10.5	13.5	18.7	23.2	<.0001
Diabetes (%)	2.5	9.3	1.3	8.1	0.01
Impaired fasting glucose‡ (%)	25.6	40.2	41.4	46.5	0.86
Hypertension (%)	15.1	35.8	26.7	40.4	0.145
Elevated triglycerides§ (%)	26.9	62.1	53.1	58.1	0.19
Low HDL Cholesterol‖ (%)	20.2	54.5	35.7	42.5	0.01
Metabolic syndrome (%)	12.9	59.7	37.5	67.3	0.002
CVD (%)	2.2	4.6	0.7	3.1	0.08
Smoking status (%)					
Current	12.8	8.3	18.8	21.2	0.04
Former	23.9	27.3	27.3	28.6	0.99
Never	63.3	64.3	54.1	50.2	0.13

VAT indicates visceral adipose tissue; TAT, thoracic periaortic fat; BMI indicates body mass index; CVD, cardiovascular disease; HDL, high-density lipoprotein.

*TAT and VAT categories are defined as high if ≥90th percentile sex-specific cut points in healthy referent sample. In the Third Generation, high TAT corresponded to volumes ≥8.79 and 15.72 cm^3 in women and men, respectively. High VAT has been previously defined as ≥1359 cm^3 in women and ≥2323 cm^3 in men.

†Sample sizes vary from row to row, as available data were used for a given characteristic.

‡Defined as fasting plasma glucose 100 to 125 mg/dL and not currently taking diabetes medication.

§Defined as ≥150 mg/dL or lipid treatment.

‖Defined as <40 mg/dL (men) and <50 mg/dL (women).

Strengths and Limitations

The major strength of our study is the relatively large sample size and community-based nature of the cohort. This allowed exploration of differences within subgroups of TAT and VAT and limited referral bias. We assessed TAT and VAT using a highly reproducible CT volumetric assessment. Certain limitations warrant discussion. The cross-sectional and observational design of the analysis prevents inferences of causality or temporality. The Framingham Heart Study is predominantly white, and results cannot be generalized to other ethnic groups. We did not have information on the severity of CVD. Our analyses were exploratory, and we did not account for multiple testing. Replication of our findings in independent cohorts is warranted. There were temporal differences between the MDCT scans and the risk factor assessments among the Offspring cohort. However, our sensitivity analysis limited to the Third-Generation cohort (in which MDCT scans and risk factor assessment occurred during the same period) demonstrated overall similar findings. Finally, our findings do not suggest that TAT quantification should be used as a clinical tool.

Conclusions

Elevated TAT is prevalent in the community and is associated with adverse cardiometabolic features, including CVD, smoking, and low HDL among individuals with normal VAT. Further work to better understand the biology of TAT and its association with metabolic and cardiovascular disease may provide insight into a potentially unique pathogenic role of periaortic fat.

Sources of Funding

This work was supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (contract N01-HC-25195). Dr. Britton was supported by a Research Career Development Award (K12 HL083786) from the National Heart, Lung, and Blood Institute.

Disclosures

Alison Pedley is an employee of Merck & Company. The remaining authors report no conflicts.

References

1. Poirier P, Giles TD, Bray GA, Hong Y, Stern JS, Pi-Sunyer FX, Eckel RH. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism. *Circulation*. 2006;113:898–918.
2. Canoy D, Boekholdt SM, Wareham N, Luben R, Welch A, Bingham S, Buchan I, Day N, Khaw K-T. Body fat distribution and risk of coronary heart disease in men and women in the European Prospective Investigation into Cancer and Nutrition in Norfolk cohort: a population-based prospective study. *Circulation*. 2007;116:2933–2943.
3. Montani JP, Carroll JF, Dwyer TM, Antic V, Yang Z, Dulloo AG. Ectopic fat storage in heart, blood vessels and kidneys in the pathogenesis of cardiovascular diseases. *Int J Obes Relat Metab Disord*. 2004;28:S58–S65.
4. Yudkin JS, Eringa E, Stehouwer CDA. "Vasocrine" signalling from perivascular fat: a mechanism linking insulin resistance to vascular disease. *Lancet*. 2005;365:1817–1820.
5. Greenstein AS, Khavandi K, Withers SB, Sonoyama K, Clancy O, Jeziorska M, Laing I, Yates AP, Pemberton PW, Malik RA, Heagerty AM. Local inflammation and hypoxia abolish the protective anticontractile properties of perivascular fat in obese patients. *Circulation*. 2009;119:1661–1670.
6. Soltis E, Cassis L. Influence of perivascular adipose tissue on rat aortic smooth muscle responsiveness. *Clin Exp Hypertens A*. 1991;13:277–296.
7. Eringa EC, Bakker W, Smulders YM, Serné EH, Yudkin JS, Stehouwer CDA. Regulation of vascular function and insulin sensitivity by adipose tissue: focus on perivascular adipose tissue. *Microcirculation*. 2007;14:389–402.
8. Chatterjee TK, Stoll LL, Denning GM, Harrelson A, Blomkalns AL, Idelman G, Rothenberg FG, Neltner B, Romig-Martin SA, Dickson EW, Rudich S, Weintraub NL. Proinflammatory phenotype of perivascular adipocytes: influence of high-fat feeding. *Circ Res*. 2009;104:541–549.
9. Schlett CL, Massaro JM, Lehman SJ, Bamberg F, O'Donnell CJ, Fox CS, Hoffmann U. Novel measurements of periaortic adipose tissue in comparison to anthropometric measures of obesity, and abdominal adipose tissue. *Int J Obes*. 2009;33:226–232.
10. Fox CS, Massaro JM, Schlett CL, Lehman SJ, Meigs JB, O'Donnell CJ, Hoffmann U, Murabito JM. Peri-aortic fat deposition is associated with peripheral arterial disease: the Framingham Heart Study. *Circ Cardiovasc Imaging*. 2010;3:515–519.
11. Lehman SJ, Massaro JM, Schlett CL, O'Donnell CJ, Hoffmann U, Fox CS. Peri-aortic fat, cardiovascular disease risk factors, and aortic calcification: the Framingham Heart Study. *Atherosclerosis*. 2010;210:656–661.
12. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C-Y, Vasan RS, Murabito JM, Meigs JB, Cupples LA, D'Agostino RB Sr, O'Donnell CJ. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. *Circulation*. 2007;116:39–48.
13. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. *Am J Epidemiol*. 1979;110:281–290.
14. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB, Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasan RS, Wolf PA, Levy D. The third generation cohort of the national heart, lung, and blood institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol*. 2007;165:1328–1335.
15. Maurovich-Horvat P, Massaro J, Fox CS, Moselewski F, O'Donnell CJ, Hoffmann U. Comparison of anthropometric, area- and volume-based assessment of abdominal subcutaneous and visceral adipose tissue volumes using multi-detector computed tomography. *Int J Obes*. 2006;31:500–506.
16. Kannel WB, Sorlie P. Some health benefits of physical activity: the Framingham Study. *Arch Intern Med*. 1979;139:857–861.
17. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III). *JAMA*. 2001;285:2486–2497.
18. Pou KM, Massaro JM, Hoffmann U, Lieb K, Vasan RS, O'Donnell CJ, Fox CS. Patterns of abdominal fat distribution. *Diabetes Care*. 2009;32:481–485.
19. Thanassoulis G, Massaro JM, Hoffmann U, Mahabadi AA, Vasan RS, O'Donnell CJ, Fox CS. Prevalence, distribution, and risk factor correlates of high pericardial and intrathoracic fat depots in the Framingham Heart Study/clinical perspective. *Circ Cardiovasc Imaging*. 2010;3:559–566.
20. Ruderman NB, Schneider SH, Berchtold P. The "Metabolically-obese", normal-weight individual. *Am J Clin Nutr*. 1981;34:1617–1621.
21. Ruderman N, Chisholm D, Pi-Sunyer X, Schneider S. The metabolically obese, normal-weight individual revisited. *Diabetes*. 1998;47:699–713.

22. Molenaar EA, Massaro JM, Jacques PF, Pou KM, Ellison RC, Hoffmann U, Pencina K, Shadwick SD, Vasan RS, O'Donnell CJ, Fox CS. Association of lifestyle factors with abdominal subcutaneous and visceral adiposity. *Diabetes Care*. 2009;32:505–510.

23. Gao Y-J, Holloway AC, Z-H Zeng, Lim GE, Petrik JJ, Foster WG, Lee RMKW. Prenatal exposure to nicotine causes postnatal obesity and altered perivascular adipose tissue function. *Obesity*. 2005;13:687–692.

24. Lohn M, Dubrovska G, Lauterbach B, Luft FC, Gollasch M, Sharma AM. Periadventitial fat releases a vascular relaxing factor. *FASEB J*. 2002;16:1057–1063.

25. Brunzell J, Goldberg A, Schwartz R. Cigarette smoking and adipose tissue lipoprotein lipase. *Int J Obes*. 1980;4:101–103.

26. Andersson K, Arner P. Systemic nicotine stimulates human adipose tissue lipolysis through local cholinergic and catecholaminergic receptors. *Int J Obes Relat Metab Disord*. 2001;25:1225–1232.

Epilepsy & Behavior 23 (2012) 57–63

Contents lists available at SciVerse ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Validation of self-reported epilepsy for purposes of community surveillance

Daniel R. Brooks [a], Ruzan Avetisyan [b], Kelli M. Jarrett [b], Amresh Hanchate [c], G.D. Shapiro [b], M.J. Pugh [d], Dan R. Berlowitz [b], David Thurman [e], Georgia Montouris [f], Lewis E. Kazis [b,*]

[a] Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA
[b] Center for the Assessment of Pharmaceutical Practices (CAPP), Department of Health Policy and Management, Boston University School of Public Health, Boston, MA, USA
[c] Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA
[d] Department of Medicine, Geriatrics and Gerontology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
[e] National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA, USA
[f] Department of Neurology, Boston University School of Medicine, Boston, MA, USA

ARTICLE INFO

Article history:
Received 28 June 2011
Revised 28 October 2011
Accepted 3 November 2011
Available online 20 December 2011

Keywords:
Epilepsy
Seizure disorder
Questionnaires
Validity
Surveillance
Self-report

ABSTRACT

We evaluated the validity of questions designed to identify lifetime and active epilepsy, medication use, and seizure occurrence on population-based surveys. Subjects were interviewed by telephone, and responses were compared with information in their medical records. Prevalence, sensitivity, specificity, and positive predictive value (PPV) were calculated. The prevalence of ever having been diagnosed with epilepsy was 3.1% by self-report and 2.7% by medical record review. Sensitivity was 84.2%, specificity was 99.2%, and PPV was 73.5% for self-reported lifetime epilepsy, and values were similar for active epilepsy. By comparison, sensitivity was higher and specificity was lower for epilepsy medication use and seizure occurrence. The PPV for seizure occurrence was substantially higher for a recall period of 12 months than for 3 months. These results compare favorably with results for other chronic conditions, such as diabetes and arthritis, and indicate that questionnaires can be used to identify epilepsy at a population level.

1. Introduction

Epilepsy is one of the most common neurological disorders in the United States [1,2]. Depending on the region and methods of measurement, recent prevalence estimates range from 0.6 to 2.0% for lifetime epilepsy and from 0.5 to 1.0% for active epilepsy [3-8].

Epilepsy has important effects on physical activities and emotional health and can lead to diminished social functioning and decrements in health-related quality of life [5,9,10]. Persons with epilepsy are more likely to report fair or poor health status and more unhealthy days than persons without epilepsy [5,10], and persons with active epilepsy report worse health than those with inactive epilepsy. Adults with epilepsy are also more likely to report being depressed than those without epilepsy, and are significantly more likely to smoke, not exercise, and be obese than people without epilepsy [3,11]. Given the significance of this condition, it is important to conduct population-based surveillance of epilepsy to measure its prevalence and monitor changes over time, identify variability in population subgroups, and provide information for purposes of planning and targeting interventions for those with epilepsy.

Various methods have been used to identify patients with epilepsy in prevalence studies. Hauser et al. used a database with detailed medical information from their large epidemiological study of Rochester, MN, USA, to identify patients with epilepsy [12]. Others have used individual practices' medical records to identify epilepsy cases [13]. Although both of these methods provide accurate epilepsy identification, they are not feasible for assessing epilepsy in larger community-wide populations, nor are they feasible for continuous monitoring.

Measurement of epilepsy prevalence in a community population can be accomplished through the use of surveys. The Epilepsy Group at the Centers for Disease Control and Prevention (CDC) developed a set of five questions (Table 1) to assess the prevalence and impact of epilepsy. Prevalence of epilepsy is determined by the first three questions in the set. Although these questions have been used in numerous surveys, including the Behavioral Risk Factor Surveillance Survey (BRFSS), the California Health Interview Survey (CHIS), and the Healthstyles Survey [3,5-8,10,11,14], they have not previously been subject to validation. The present study sought to validate the questions used to estimate epilepsy prevalence by comparing patients' responses with documentation in their medical records.

* Corresponding author at: Center for the Assessment of Pharmaceutical Practices (CAPP), Department of Health Policy and Management, Boston University School of Public Health, Boston, MA 02118, USA.
E-mail address: lek@bu.edu (L.E. Kazis).

1525-5050/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.yebeh.2011.11.002

Table 1
Epilepsy screening questions.

1. Have you ever been told by a doctor that you have a seizure disorder or epilepsy?
2. Are you currently taking any medicine to control your seizure disorder or epilepsy?
3. How many seizures of any type have you had in the last 3 months?
4. In the past year, have you seen a neurologist or epilepsy specialist for your epilepsy or seizure disorder?
5. During the past month, to what extent has epilepsy or its treatment interfered with your normal activities like working, school, or socializing with family or friends?

2. Methods

2.1. Study population

The study was conducted among a sample of patients at Boston Medical Center (BMC) in Boston, MA, USA, who were ≥18 years old, able to speak English, and cognitively able to participate in a telephone interview. BMC is an urban multispecialty teaching hospital that provides primary and specialty care and is also a major source of care for low-income residents and immigrants. The study was approved by the Boston University Medical Center Institutional Review Board.

Because a simple random sample of patients would yield few cases of epilepsy, we sampled patients after stratifying according to probability of having an epilepsy diagnosis. To identify putative epilepsy cases, we searched the electronic administrative/billing and medical record system, which covers visits back to 2000, for all patients with an ICD-9 code of either 345.XX[1] or 780.39 or a listing of the terms "epilepsy" or "seizure disorder" in the medical record problem list. Similar methodology has been used by others to identify epilepsy cases using ICD-9 codes [15-18]. For putative noncases, we started with a list of all patients who had a doctor in the primary care clinic and who met the general study eligibility criteria, and removed persons who met the criteria for putative cases.

We narrowed the sampling frame for both putative cases and noncases to patients who had made either: (1) two or more visits to the primary care or neurology clinics within the past 2 years, or (2) one visit to the primary care clinic or neurology clinic and at least one nonemergency inpatient visit within the past 2 years. This criterion was imposed to increase the likelihood that there would be a current telephone number for contacting subjects as well as sufficient information in the medical record to establish the validity of their survey responses. Because of these concerns, subjects whose only contact with BMC was through the emergency department were excluded. Although the emergency department is a potentially important point of care for persons with epilepsy, a preliminary review of medical records found that information documented in the record of an emergency department visit was insufficient to identify a case of epilepsy unless there were also other visits to BMC. A random sample of eligible putative cases and noncases were then selected to be contacted for a phone interview.

2.2. Data collection

Patient information was obtained from different sources; those used in this analysis include the administrative/billing database, telephone interviews, and medical chart reviews. The administrative/billing database provided basic demographics: sex, age, race/ethnicity, marital status, zip code of residence (which together with 2000 Census data was used to assign the median neighborhood income to each subject), medical insurance information, and health care utilization data.

2.3. Patient recruitment and telephone interviews

Interviews were conducted over a 10-month period from December 2006 to October 2007. After receiving permission from their neurologist or primary care physician, we used telephone numbers from the administrative record to contact potential respondents. Trained interviewers made a minimum of nine attempts to reach each patient at different days and times, including evenings and weekends. For patients we were unable to contact and who had made a subsequent visit to BMC, we consulted the administrative record for an updated telephone number. On reaching a potential respondent, we described the study in a general manner that did not reveal its purpose or its focus on epilepsy. Persons who agreed to participate were consented by phone and administered the questionnaire; no proxy respondents were permitted. All interviews were conducted in English. The average length of the interview was 15 minutes.

The telephone interview questionnaire (see Appendix for entire survey instrument) contained the five Behavioral Risk Factor Surveillance System (BRFSS) screening questions regarding epilepsy (Table 1), as well as the VR-12 Health Survey instrument, BRFSS Healthy Days questions, and demographic items. We also added several supplementary epilepsy and seizure-related questions designed to provide more detail and help interpret answers to the screening questions. Questions included age at first diagnosis, causes of seizures if not epilepsy, number of seizures (lifetime and past 12 months), and site of medical care for ongoing epilepsy treatment and for most recent seizure. All five screening questions were asked of all study participants regardless of their responses to a previous item. To assess reliability of responses, we called back 50 putative cases within 4–6 weeks of their initial interview and readministered the screening questions and selected additional items.

2.4. Medical chart reviews

Data collection from the medical charts of telephone questionnaire respondents was performed by specially trained research staff with a clinical background. The chart review instrument comprised clinical variables capable of validating the screening questions and was developed based on the input of study epileptologists and neurologists. The form and protocols for the chart reviews were developed and tested in an iterative fashion to accommodate the variability in chart documentation as practiced in different settings and by different physicians. The chart review forms are available from the corresponding author (L.E.K.).

The reviewers abstracted data related to epilepsy, including its assessment, treatment, side effects, and complications, using clinician notes for primary care and specific medical subspecialty visits, as well as documented data on emergency department and inpatient visits. Chart reviews covered the period that a patient began receiving care at BMC beginning as far back as 2000, when the electronic medical record system first became operational, through the date of the phone interview. The note from each visit was reviewed and

[1] 345 is a nonspecific code for epilepsy; epilepsy diagnosis is usually documented with 345 and single- or double-digit index entries: 345.xx. 780.39 is a nonspecific code for seizures.

information on diagnoses, seizure history, referrals, and changes in medications contained in the note was entered onto the data abstraction instrument. Abstractions were aided by a glossary of medical terms, which also contained the names of all medications used in the treatment of epilepsy and described each respective class of drugs.

Specific data elements captured fell into three general categories: demographics and utilization of services, seizure history and diagnosis, and treatment and medications. The seizure history and diagnosis section included the occurrence, frequency, and recurrence of seizures and diagnoses of epilepsy or "seizure disorder." The treatment and medications section covered dosage and frequency and whether a recent change had been made in the medication. Diagnoses of other medical conditions were also collected from the chart and used to construct indices of physical and mental comorbidities [19].

A sample of records were reviewed a second time by a second abstractor to determine interrater reliability. A higher percentage of charts were tested during the first 2 months of data abstraction to identify and correct any areas of particularly high disagreement among the abstractors.

2.5. Validation of epilepsy test items

We used information collected from patient medical records at BMC as the standard against which to compare self-report data from the telephone questionnaire for screening questions Q1–Q3 (Table 1). A diagnosis of epilepsy (Q1) was confirmed based on documentation in the patients' medical record. Medical records documented the clinical histories of epilepsy, as well as EEG and other neuroimaging test results, blood work, and referrals with varying degrees of detail. Thus, we accepted physician diagnoses as evidence of epilepsy when these were documented in the medical record without ambiguity or inconsistency. However, when ambiguities were noted, an epileptologist G.M. who was an internal consultant to the study and familiar with the study protocols reviewed the record, considering all available clinical information in the charts, including test results, to make the case determination.

The criterion for a diagnosis of epilepsy was explicit mention of "epilepsy" or "seizure disorder" in either a clinician note or the patient's problem list. Current medication use (Q2) was confirmed based on prescriptions noted in the medical record at the time of the telephone interview. Occurrence of seizures in the past 3 months (Q3) was validated based on mention of a seizure at any visit to BMC during that time frame as calculated from the date of interview. We were also interested in the validity of 12-month recall of seizures and so included this item in the telephone questionnaire, which was validated as described above but for a 12-month period.

2.6. Statistical analysis

All data from administrative records, questionnaires, and chart reviews were entered into a relational database and linked by a common study ID number. We first calculated the adjusted response rate as the number of completed interviews divided by the total sample excluding those determined to be ineligible and those expected to be ineligible among those who were not contacted based on the proportion ineligible among those contacted [20,21]. We also calculated the cooperation rate (completed interviews as proportion of potential respondents reached and determined to be eligible). Response rates are reported separately for putative epilepsy cases and noncases and stratified by the demographic, utilization, and clinical factors listed in Tables 2A and 2B.

The primary measures used to assess the accuracy of Q1–Q3 were sensitivity, specificity, positive predictive value, and the difference in the estimate of prevalence based on self-report versus chart-based data. Because the sample had been constructed to oversample persons likely to have epilepsy, we weighted three strata, as identified hierarchically in the electronic billing system [putative cases based on (1) ICD-9 345.1 or (2) ICD-9 780.39 or problem list mention, and (3) putative noncases], so that their distribution would be more representative of a sample randomly drawn from a community-based population. We weighted the data to the distribution of the three strata in the electronic billing system for patients who attended the Primary Care Center at BMC, which we took as the closest approximation to a community-based sample among the BMC patient population. The distribution of the three strata was 1.76% for 345.1 putative cases, 2.44% for 780.39 putative cases, and 95.8% for putative noncases. We weighted each stratum so that it summed to its proportion of the total sample size. Weights were equal to 0.064, 0.131, and 4.46 for the three strata, respectively.

Using the weighted data, we first calculated measures of validity for reported lifetime and active epilepsy. We defined reported lifetime epilepsy as an affirmative answer to the first screening question: "Have you ever been told by a doctor that you have a seizure disorder or epilepsy?" We defined active epilepsy as a report of lifetime epilepsy along with (1) currently being on medication for epilepsy and/or (2) having had a seizure in the past 3 months [5-8]. We also tested an alternative definition in which the period of last seizure was extended to 12 months. We excluded seven participants who could not be categorized as self-reporting lifetime or active epilepsy because they responded "don't know" or "refused" to the relevant questions.

We also assessed validity of lifetime and active epilepsy stratified by demographic (sex, age, race, median income of zip code of

Table 2A
Adjusted response rates, cooperation rates, and number and distribution of interviewees, by demographic and health-related characteristics, among putative cases.

	Adjusted response rate	Cooperation rate	Interviewed	Percent of total interviewed
Total	38%	57%	467	100%
Gender				
Female	40%	57%	262	56%
Male	36%	57%	205	44%
Age				
18-34	33%	56%	100	21%
35-64	39%	60%	301	64%
≥65	38%	49%	66	14%
Race				
Black	37%	56%	200	43%
Hispanic	39%	59%	30	6%
White	39%	58%	207	44%
Other	40%	61%	30	6%
Income tertile				
1 (lowest)	39%	59%	166	36%
2 (middle)	36%	58%	128	27%
3 (highest)	39%	56%	173	37%
Time from visit before initial contact date				
<1 mo	48%	60%	216	46%
1-6 mo	38%	57%	184	39%
6-12 mo	24%	47%	35	7%
1-2 yrs	25%	57%	32	7%
Total neurology + PCP visits in last 2 years				
1-3	34%	56%	169	36%
4-6	35%	56%	113	24%
7-9	43%	58%	73	16%
≥10	47%	60%	112	24%
Physical comorbidities				
0-1	33%	54%	167	36%
2-3	41%	63%	129	28%
4-5	40%	57%	76	16%
> 5	43%	57%	95	20%
Mental comorbidities				
0	37%	55%	244	52%
1	40%	58%	118	25%
2-5	38%	63%	105	22%

Note. PCP, primary care physician.

Table 2B
Adjusted response rates, cooperation rates, and number and distribution of interviewees, by demographic and health-related characteristics, among putative noncases.

	Adjusted response rate	Cooperation rate	Interviewed	Percent of total interviewed
Total	28%	43%	216	100%
Gender				
Female	30%	44%	125	58%
Male	25%	41%	91	42%
Age				
18-34	22%	43%	28	13%
35-64	29%	46%	146	68%
≥65	27%	35%	42	19%
Race				
Black	25%	40%	115	53%
Hispanic	27%	39%	13	6%
White	34%	49%	66	31%
Other	30%	45%	22	10%
Income tertile				
1 (lowest)	23%	37%	71	33%
2 (middle)	31%	47%	68	31%
3 (highest)	31%	47%	77	36%
Time from visit before initial contact date				
< 1 mo	33%	43%	112	52%
1-6 mo	28%	42%	81	38%
6-12 mo	23%	51%	20	9%
1-2 yrs	7%	27%	3	1%
Total neurology + PCP visits in last 2 years				
1-3	18%	35%	47	22%
4-6	31%	50%	70	32%
7-9	33%	45%	54	25%
≥10	33%	42%	45	21%
Physical comorbidities				
0-1	22%	43%	42	19%
2-3	27%	45%	68	31%
4-5	27%	40%	44	20%
> 5	35%	43%	62	29%
Mental comorbidities				
0	23%	35%	102	47%
1	33%	47%	58	27%
2-5	37%	60%	56	26%

Note. PCP, primary care physician.

residence), utilization (most recent visit before initial phone contact, total number of neurology and primary care visits in past 2 years), and clinical (physical health comorbidities, mental health comorbidities) variables. We summed the current physical and mental health comorbidities based on ICD-9 codes listed in the administrative/billing database from 30 physical health and 6 mental health conditions that constitute the Selim Comorbidity Index, a previously validated measure [19]. However, we excluded seizures, which is one of the conditions in the original index.

One analytical complication in the stratified analysis arose from the presence of a single false-positive report of epilepsy among the 215 subjects in the putative noncase stratum, all of whom were determined not to have epilepsy by chart review. Because subjects in this stratum had such large weights relative to the two putative case strata, the demographic or clinical category that happened to contain this individual always had lower accuracy than other strata. Because of its distorting effect on the results, we excluded this one observation from the analysis of stratum-specific measures.

We used the same accuracy measures to evaluate the performance of Q2 and Q3. We restricted analyses to respondents with self-reported lifetime epilepsy, regardless of confirmation by chart review, as this represented the performance of the questions as they would function in an actual survey setting where all respondents who answered affirmatively to the first question would be asked the remaining questions. We again excluded the single putative noncase with a false-positive report of lifetime epilepsy. This individual constituted the only observation in this stratum and greatly inflated the standard error while providing essentially no information about the true distribution of responses. Therefore, the analysis included only the two putative case strata, which together accounted for 86% of the weighted value of all participants with self-reported epilepsy.

For all of the above analyses, we used the SVYSET procedure in STATA 10.1, accounting for different weights and the stratified sampling scheme [22]. Because the standard symmetric confidence intervals based on the normal approximation to the binomial are increasingly inaccurate at proportions very close to 0 or 1, we calculated asymmetric 95% confidence intervals based directly on the binomial distribution [23].

3. Results

We sampled a total of 1727 putative epilepsy cases (1033 ICD-9 345.xx, 694 ICD-9 780.39) and 1100 putative noncases. Compared with the latter, putative cases were more likely to be white (38% vs 19%), have made fewer visits to BMC in the past 2 years (one to three visits: 43% vs 30%), and have fewer physical health comorbidities (none or one comorbidity: 41% vs 22%). Among the putative epilepsy cases, 332 were ineligible, 580 could not be reached, 348 refused to be interviewed, and 467 completed the telephone interview. The corresponding figures among putative noncases were 209 ineligible, 387 unable to be reached, 288 refused, and 216 completed.

Response rates were 38% among putative cases and 28% among putative noncases, whereas the cooperation rates were 57 and 43%, respectively (Table 2A). Response and cooperation rates were higher for putative cases than noncases in virtually all demographic, utilization, and clinical categories. Factors associated with higher response rates among both putative cases and noncases included more recent visit to BMC, greater number of visits in past 2 years, and more physical health comorbidities. Among putative noncases, response rates were also higher among persons with a larger number of mental health comorbidities. Differences in response rates between putative cases and noncases were particularly large among those who were residents of the lowest income tertile of zip codes (39% vs 23%), had not been to BMC for >1 year (25% vs 7%), and had made 1–3 (34% vs 18%) or ≥10 (47% vs. 32%) visits to BMC in the past 2 years.

3.1. Lifetime and active epilepsy

The results for the validation of lifetime and active epilepsy are summarized in Table 3. The prevalence of lifetime epilepsy in this sample was 3.1% by self-report and 2.7% by chart review. Sensitivity and specificity of self-reported lifetime epilepsy were 84.2 and 99.2%, respectively, and the positive predictive value (PPV) was 73.5%. Approximately 85% of subjects with lifetime epilepsy had active epilepsy by both chart review and self-report. Of the 302 patients with active epilepsy based on chart review, only one was

Table 3
Validation measures for self-report of lifetime and active epilepsy.

	% (95% CI)	
	Lifetime[a]	Active[b]
N	675	675
Prevalence		
Self-report	3.1 (2.3–4.1)	2.6 (1.8–3.6)
Chart review	2.7 (2.5–2.9)	2.3 (2.2–2.5)
Sensitivity	84.2 (79.3–88.1)	81.1 (75.8–85.5)
Specificity	99.2 (97.5–99.7)	99.3 (97.5–99.8)
Predictive value		
Positive	73.5 (48.1–89.2)	73.9 (43.3–91.3)
Negative	99.6 (99.4–99.7)	99.5 (99.4–99.7)

[a] Lifetime epilepsy: reported ever being told by a doctor that had epilepsy or a seizure disorder.
[b] Active epilepsy: lifetime epilepsy plus reported currently taking medicine for epilepsy or a seizure disorder OR reported having a seizure within the past 3 months.

not taking medication but had had a seizure in the 3 months prior to the interview. Sensitivity, specificity, predictive value, and the difference in active epilepsy prevalence between self-report and chart review were almost the same as for lifetime epilepsy. Extending the time since last seizure to 12 months had almost no impact on the prevalence or accuracy of active epilepsy, because the great majority of respondents qualified on the basis of taking medication.

Table 4 outlines the results of a stratified analysis of the accuracy of active epilepsy based on demographic and clinical variables. Prevalence measured by self-report was generally lower than that measured by chart review, but differences were generally small (range: 0.1–0.8%). Measures of specificity also varied within a small range (99.5–100%). There were no significant differences in sensitivity, but rates were lower for older adults and those who had not visited BMC for ≥6 months. Persons with mental comorbidities had a significantly lower PPV than those without any comorbidities.

We also used logistic regression analysis to simultaneously adjust for all the characteristics in Table 4 to identify the primary predictors of a false (positive or negative) response for active epilepsy. The odds of a false response were higher among subjects who were male, had more physical or mental comorbidities, and had not visited BMC for ≥6 months (data not shown).

We also conducted a supplementary analysis to explore reasons for the false-positive responses for lifetime epilepsy. Because Q1 asks "Have you ever been told by a doctor that you have a *seizure disorder* or *epilepsy*," we hypothesized that false-positive responses were more likely to occur among subjects who had experienced a seizure which had not been diagnosed with epilepsy. A subsequent question in the telephone interview asked whether patients were specifically diagnosed as having epilepsy or another seizure disorder, and 62% of false-positive responses named some sort of seizure disorder but not explicitly epilepsy, compared with 41% of true-positive responses.

The medical history of the single false-positive case from the putative noncase stratum illustrates the potential for this problem. In addition to responding affirmatively to the question about lifetime epilepsy, the subject also reported currently using medication for epilepsy or seizure disorder. In response to follow-up questions, the subject could not say whether the diagnosis was epilepsy or something else. Review of the medical record revealed that the subject had suffered serious head trauma several years ago. The subject was started on an antiseizure medication (phenytoin), which appears to have been discontinued. The subject reports never having any seizures in two separate sections of the questionnaire, and there is no evidence of seizures in the medical record. On the basis of these facts, we suggest that the subject likely was told of the risk of seizure following the head trauma and therefore interpreted the term *seizure disorder* in the questionnaire as applicable. The subject continues to be followed for chronic headaches, which are likely linked in the subject's mind to the underlying "seizure disorder" resulting from the head trauma and, therefore, may explain the affirmative answer to current use of medication for a seizure disorder.

3.2. Medication use and seizure occurrence

The prevalence of current medication use among those who self-reported having epilepsy was similar by self-report and chart review (78.3% vs 81.8%) (Table 5). Sensitivity was 90% and specificity 74%. As expected given the high prevalence, the PPV (93.9%) was higher than the negative predictive value (62.1%). The sensitivity

Table 4
Validation measures for active epilepsy by demographic and health-related characteristics.

	N[a]	(95% CI)					
		Prevalence		Sensitivity	Specificity	Predicitive Value	
		Self-Report	Chart-Review			Positive	Negative
Gender							
Female	382	2.2 (2.0-2.4)	2.3 (2.1-2.5)	86.1 (79.3-91.0)	99.8 (99.6-99.9)	90.6 (83.9-94.7)	99.7 (99.5-99.8)
Male	292	2.1 (1.8-2.4)	2.5 (2.2-2.8)	74.8 (65.6-82.1)	99.7 (99.5-99.9)	87.4 (78.1-93.1)	99.4 (99.1-99.6)
Age							
18-34	126	3.2 (2.7-3.8)	4.0 (3.4-4.6)	78.4 (66.2-87.0)	99.9 (99.6-99.9)	96.9 (88.3-99.2)	99.1 (98.5-99.5)
35-64	439	2.3 (2.0-2.5)	2.3 (2.1-2.5)	84.1 (77.5-89.1)	99.7 (99.5-99.8)	86.3 (79.5-91.1)	99.6 (99.5-99.7)
≥65	109	1.0 (0.8-1.3)	1.4 (1.1-1.8)	69.7 (51.8-83.1)	100 (99.8-100)	96.8 (79.9-99.6)	99.6 (99.1-99.8)
Race							
White	263	3.3 (2.9-3.7)	3.8 (3.4-4.3)	78.3 (69.9-84.9)	99.7 (99.4-99.9)	91.3 (83.4-95.7)	99.1 (98.7-99.4)
Black	306	1.6 (1.4-1.8)	1.7 (1.5-1.9)	82.2 (73.3-88.6)	99.8 (99.7-99.9)	88.5 (79.6-93.8)	99.7 (99.5-99.8)
Other	104	1.9 (1.5-2.3)	1.8 (1.5-2.2)	88.9 (71.5-96.2)	99.7 (99.3-99.9)	85.0 (66.8-94.1)	99.8 (99.4-99.9)
Income tertile							
Low	235	2.1 (1.8-2.5)	2.2 (2.0-2.6)	82.5 (72.7-89.4)	99.7 (99.5-99.9)	87.1 (76.8-93.2)	99.6 (99.3-99.8)
Middle	193	1.9 (1.7-2.2)	2.0 (1.7-2.3)	89.3 (79.2-94.9)	99.8 (99.6-99.9)	91.3 (81.1-96.2)	99.8 (99.6-99.9)
High	246	2.3 (2.0-2.7)	2.8 (2.5-3.1)	75.1 (65.7-82.6)	99.7 (99.5-99.9)	89.8 (80.2-95.0)	99.3 (99.0-99.5)
Physical comorbidities							
0-1	173	3.2 (2.8-3.7)	3.6 (3.2-4.1)	86.2 (76.5-92.3)	99.9 (99.5-100)	96.9 (87.5-99.3)	99.5 (99.1-99.7)
2-3	185	1.8 (1.5-2.1)	1.9 (1.6-2.2)	83.2 (72.1-90.5)	99.8 (99.5-99.9)	87.1 (75.3-93.8)	99.7 (99.4-99.8)
4-5	140	2.2 (1.8-2.7)	2.7 (2.2-3.1)	71.3 (58.1-81.7)	99.6 (99.3-99.8)	84.4 (70.3-92.6)	99.2 (98.7-99.5)
> 5	176	1.7 (1.4-2.1)	1.8 (1.5-2.1)	82.3 (69.6-90.4)	99.8 (99.5-99.9)	86.6 (72.9-93.9)	99.7 (99.4-99.8)
Mental comorbidities							
0	343	2.4 (2.2-2.7)	2.9 (2.6-3.1)	82.7 (75.6-88.1)	99.9 (99.9-100)	97.7 (94.1-99.2)	99.5 (99.3-99.7)
1	173	2.1 (1.8-2.5)	2.0 (1.7-2.4)	84.6 (72.3-92.1)	99.6 (99.3-99.8)	80.8 (68.2-89.2)	99.7 (99.4-99.8)
2-5	158	1.6 (1.3-2.0)	1.7 (1.4-2.1)	71.9 (57.8-82.8)	99.6 (99.3-99.8)	77.6 (62.1-88.0)	99.5 (99.1-99.7)
Number of visits in last 2 years							
1-3	213	3.5 (3.0-4.1)	4.0 (3.5-4.5)	79.2 (69.4-86.5)	99.6 (99.3-99.8)	89.9 (80.5-95.0)	99.1 (98.7-99.5)
4-6	180	1.7 (1.5-2.0)	2.0 (1.7-2.2)	81.2 (70.0-88.9)	99.9 (99.7-100)	94.0 (84.1-97.9)	99.6 (99.3-99.8)
7-9	126	1.3 (1.1-1.7)	1.6 (1.3-1.9)	79.9 (64.4-89.8)	99.9 (99.7-100)	94.1 (77.8-98.6)	99.7 (99.4-99.8)
≥10	155	2.3 (1.9-2.8)	2.2 (1.8-2.6)	85.9 (73.2-93.1)	99.5 (99.1-99.7)	80.1 (66.4-89.1)	99.7 (99.4-99.9)
Time from visit before initial contact							
< 1 mo	322	1.9 (1.7-2.2)	2.0 (1.8-2.3)	84.3 (76.0-90.1)	99.7 (99.6-99.9)	87.6 (79.2-92.9)	99.7 (99.5-99.8)
1-6 mo	263	2.2 (1.9-2.5)	2.5 (2.2-2.8)	84.2 (75.8-90.1)	99.9 (99.7-100)	95.3 (87.7-98.3)	99.6 (99.3-99.7)
6-12 mo	54	1.7 (1.2-2.4)	2.3 (1.8-3.0)	64.4 (41.9-82.0)	99.8 (99.0-99.9)	86.9 (57.6-97.0)	99.2 (98.4-99.6)
≥12 mo	35	9.0 (6.3-12.7)	9.8 (6.9-13.6)	63.8 (41.1-81.7)	96.9 (92.5-98.8)	69.4 (42.9-87.2)	96.1 (91.5-98.3)

Table 5
Validation measures for current medication use and time since last seizure (3 months and 12 months) among persons reporting lifetime Epilepsy[a].

	Q2 Medication Use	Q3 Seizure	
		3 Months	12 Months
N	321	321	321
Prevalence			
Self Report % (95% CI)	78.3 (73.0-82.8)	35.1 (29.9-40.7)	51.9 (46.0-57.8)
Chart Review % (95% CI)	81.8 (76.7-86.0)	19.3 (15.5-23.8)	39.2 (33.7-44.9)
Sensitivity % (95% CI)	89.9 (85.1-93.3)	90.7 (79.0-96.2)	92.1 (85.3-95.9)
Specificity % (95% CI)	74.0 (58.4-85.2)	78.1 (72.1-83.1)	73.9 (66.4-80.2)
Predictive Value			
Positive % (95% CI)	93.9 (89.5-96.6)	49.5 (40.4-58.8)	69.4 (61.3-76.5)
Negative % (95% CI)	62.1 (48.1-74.3)	97.3 (93.2-98.9)	93.5 (88.0-96.6)

[a] Excludes single false-positive putative non-case (see Methods).

and specificity of report of seizure occurrence in the past 3 months were very similar to those for medication use. However, there was a substantial difference in seizure prevalence between self-report and chart review (35.1% vs 19.3%). The PPV was 49.5%. When seizure occurrence in the past 12 months was substituted for the 3-month recall period, the prevalence of both self-reported and chart-verified seizure increased substantially to 51.9 and 39.2%, respectively, although there was little change in the extent of the difference between the two measures (15.8% at 3 months and 12.7% at 12 months). Other measures of validity were also similar with the exception of the PPV, which increased from 49.5% for 3-month to 69.4% for 12-month recall.

It is possible that the low PPV and substantial disparity between self-reported and chart-based prevalence of seizure were due not to a problem with the question or subjects' memory, but rather to the fact that we only had access to BMC medical records. If an individual had a seizure that was not treated or was treated elsewhere and not documented in the BMC medical record, the response would have been classified as a false positive. Subjects who were classified as making true-positive responses were more likely than those classified as making false-positive responses to report having come to BMC for treatment of their last seizure (66% vs 35%). To better understand the possible impact of treatment location, we reanalyzed the data for 3-month recall based on the assumption that the ratio of true to false positives among those who did not go to BMC for care of their last seizure would have been the same as for those who did get care at BMC. The difference between self-reported and chart-based prevalence decreased from 15.8 to 10.4%, whereas sensitivity increased by 2.9%, specificity by 8.7%, and the PPV by 17.4%.

Among the 50 subjects who were interviewed a second time to assess test–retest reliability, κ values were 0.64 for lifetime epilepsy, 0.88 for medication use, and 0.65 for seizure occurrence.

4. Discussion

Well-validated epilepsy screening assessments are important for purposes of monitoring both the burden of epilepsy and the needs of people with epilepsy at a community level. Tools for surveillance systems are important for planning health care delivery systems, designing appropriate clinical care services, and designing and implementing preventive services for policy makers interested in future resource allocations. The results of this study indicate that the questions developed by the CDC Epilepsy Group are valid for the identification of epilepsy. Sensitivity and specificity exceeded 80 and 99%, respectively, for both lifetime and active epilepsy. In addition, prevalence as measured by self-report was within 0.4% of the estimate based on chart review prevalence for lifetime and active epilepsy. The absolute estimates of self-report prevalence, specificity, and PPV were higher in the stratified results than in the overall analysis, which was due to the exclusion of the single false-positive putative noncase who had a highly disproportionate effect on the estimates. However, the focus in the stratified analysis should be on the relative variation in estimates between strata rather than the absolute level. Few meaningful differences were seen for any measure by stratum of demographic, clinical, or medical care utilization characteristics. However, being male, having more physical or mental comorbidities, and not having visited BMC for $\geq$6 months were associated with inaccurate responses.

The study used a stratified sampling design which greatly overrepresented people with epilepsy. Had we used a simple random sample, it would have been necessary to interview more than 10,000 subjects to identify the same number of people who reported having epilepsy. We weighted our data to reflect the prevalence of putative epilepsy among primary care patients as a proxy for a community-based sample. It was important to weight the sample in this manner because the PPV of self-reported epilepsy is affected by the prevalence in the population in addition to the accuracy of the test. Had the data not been weighted, the estimate of the PPV would have been artificially high. Even so, probably because the sample was drawn from a medical setting, the prevalence of epilepsy—both self-report and chart-based—was somewhat higher in our study than in community-based estimates, particularly for active epilepsy.

In our study, we asked all the screening questions of each subject regardless of how they responded to the first question concerning whether they had ever been told they had epilepsy or a seizure disorder. In a typical setting, however, only persons responding affirmatively to this question would proceed to the subsequent questions. The PPV provides information regarding the universe of people being asked the remaining questions and the proportion who actually have epilepsy. Both lifetime and active epilepsy had a PPV of 74%, although the estimate was somewhat imprecise (as reflected by the wide confidence intervals) owing to the influence of the false-positive putative noncase. Based on this estimate, it appears that approximately one of four people answering the remaining questions does not actually have epilepsy.

The levels of sensitivity and specificity in this study are similar or higher than have been reported in studies that assessed the validity of self-report compared with medical records for other chronic conditions. For example, a study of 599 HMO subscribers in Colorado compared telephone self-report for diabetes with medical records and found 73% sensitivity and 99.3% specificity [24]. A similar study among $\geq$45-year-old residents of Olmsted County, MN, USA, that used self-administered questionnaires showed 66% sensitivity and 99.7% specificity for diabetes [25]. A study that compared telephone self-report of arthritis with rheumatologist examination found a sensitivity of 70.8% and specificity of 70.3% [26]. Diabetes and arthritis exhibit similarities to epilepsy in that all three conditions are long-lasting and share similar impacts on quality of life including physical, psychological, and social functioning.

There are also some limitations to this study. Despite multiple efforts to contact and enroll subjects, the response and cooperation rates were low, as might be expected among a disproportionately low-income and transient population. Despite the fact that we did not mention that the study concerned epilepsy, the rates were higher for persons with putative diagnoses of epilepsy. To the extent that subjects in the study differed in the accuracy of their responses compared with those who did not participate, our results could have been affected. However, as previously noted the results of this study are comparable to those reported in other studies assessing the validity of self-report compared with medical records. The biggest differences between responders and nonresponders was that those who responded were more likely to have more contact, and more recent contact, with BMC and were also more likely to have more physical and mental comorbidities. More recent contact was also associated with greater accuracy, which would suggest that the

observed measures of sensitivity and specificity may be somewhat overstated. However, this effect appears to be modest because only 13% of subjects had not had recent contact.

There are also limitations to using chart review as the validation standard. For patients whose main source of epilepsy care is elsewhere, there may not be sufficient documentation in the BMC record to validate the screening questions and the occurrence of seizures, as suggested by the the fact that subjects categorized as false positive were more likely to have reported receiving care for the most recent seizure elsewhere or not at all. Additionally, the independence of charts from patient self-report may be overestimated. The source of documentation of a seizure occurrence in a chart may be the patient's self-report to the physician, which can make seizures particularly problematic to validate using medical records, unlike other conditions such as hypertension and hypercholesterolemia that can be measured using objective tests in a clinical setting. Even the presence of lifetime epilepsy in the medical record may rely on the patient's self-report, particularly if he or she is an inactive case with no current seizures or medication use.

Finally, our sampling frame in general may not be generalizable to a community sample or sample with different sociodemographic characteristics. We worked with a population receiving care in an urban tertiary care setting and, in particular, one of the only safety-net hospitals in the area that provides free care. Therefore, subjects included in this study are likely sicker and more disabled, as well as poorer and less likely to have health insurance, than a general community-based sample. Thus, it would be useful to replicate this study in other populations to assess the generalizability of these findings.

In summary, for population surveillance of epilepsy prevalence, as well as medication use and seizure occurrence, future population-based assessments using these items can provide important information regarding the prevalence and characteristics of epilepsy in the US population.

Acknowledgments

This publication was supported by Cooperative Agreement 5 U48 DP00058 from the Centers for Disease Control and Prevention. The findings and conclusions are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.yebeh.2011.11.002.

References

[1] Sander JW, Shorvon SD. Epidemiology of the epilepsies. J Neurol Neurosurg Psychiatry 1996;61:433–43.
[2] Hirtz D, Thurman DJ, Gwinn-Hardy K, Mohammed M, Chaudhuri AR, Zalutsky R. How common are the "common" neurological disorders? Neurology 2007;68:326–37.
[3] Kobau R, Dilorio CA, Price PH, et al. Prevalence of epilepsy and health status of adults with epilepsy in Georgia and Tennessee: Behavioral Risk Factor Surveillance System, 2002. Epilepsy Behav 2004;5:358–66.
[4] Kelvin EA, Hesdorffer DC, Bagiella E, et al. Prevalence of self-reported epilepsy in a multiracial and multiethnic community in New York City. Epilepsy Res 2007;77:141–50.
[5] Kobau R, Zahran H, Grant D, Thurman DJ, Price PH, Zack MM. Prevalence of active epilepsy and health-related quality of life among adults with self-reported epilepsy in California: California Health Interview Survey, 2003. Epilepsia 2007;48:1904–13.
[6] Elliott JO, Moore JL, Lu B. Health status and behavioral risk factors among persons with epilepsy in Ohio based on the 2006 Behavioral Risk Factor Surveillance System. Epilepsy Behav 2008;12:434–44.
[7] Ferguson PL, Chiprich J, Smith G, et al. Prevalence of self-reported epilepsy, health care access, and health behaviors among adults in South Carolina. Epilepsy Behav 2008;13:529–34.
[8] Kobau R, Zahran H, Thurman DJ, et al. Epilepsy surveillance among adults—19 states, Behavioral Risk Factor Surveillance System, 2005. MMWR Morb Mortal Wkly Rep 2008;57(SS06):1–20.
[9] Sample PL, Ferguson PL, Wagner JL, Pickelsimer EE, Selassie AW. Experiences of persons with epilepsy and their families as they look for medical and community care: a focus group study from South Carolina. Epilepsy Behav 2006;9:649–62.
[10] Kobau R, Gilliam F, Thurman DJ. Prevalence of self-reported epilepsy or seizure disorder and its associations with self-reported depression and anxiety: results from the 2004 Healthstyles Survey. Epilepsia 2006;47:1915–21.
[11] Elliott JO, Lu B, Shneker B, Charyton C, Moore JL. Comorbidity, health screening, and quality of life among persons with a history of epilepsy. Epilepsy Behav 2009;14:125–9.
[12] Hauser WA, Annegers JF, Kurland LT. Incidence of epilepsy and unprovoked seizures in Rochester, Minnesota: 1935–1984. Epilepsia 1993;34:453–68.
[13] Heaney DC, MacDonald BK, Everitt A, et al. Socioeconomic variation in incidence of epilepsy: prospective community based study in south east England. Br Med J 2002;325:1013–6.
[14] Centers for Disease Control and Prevention. Health-related quality of life among persons with epilepsy—Texas, 1998. MMWR Morb Mortal Wkly Rep 2001;50:24–6.
[15] Holden EW. Thanh Nguyen H, Grossman E, et al. Estimating prevalence, incidence, and disease-related mortality for patients with epilepsy in managed care organizations. Epilepsia 2005;46:311–9.
[16] Wright J, Pickard N, Whitfield A, Hakin N. A population-based study of the prevalence, clinical characteristics and effect of ethnicity in epilepsy. Seizure 2000;9:309–13.
[17] Parko K, Thurman DJ. Prevalence of epilepsy and seizures in the Navajo Nation 1998–2002. Epilepsia 2009;50:2180–5.
[18] Pugh MJ, Van Cott AC, Cramer JA, et al. Trends in antiepileptic drug prescribing for older patients with new-onset epilepsy: 2000–2004. Neurology 2008;70:2171–8.
[19] Selim AJ, Fincke G, Ren XS, et al. Comorbidity assessments based on patient report: results from the Veterans Health Study. J Ambul Care Manage 2004;27:281–95.
[20] Standard definitions: final dispositions of case codes and outcome rates for surveys. Lenexa, KS: Am. Assoc. for Public Opinion Research; 2000.
[21] Lynn P, Beerten R, Laiho J, Martin J. Recommended standard final outcome categories and standard definitions of response rate for social surveys. Colchester: Institute for Social and Economic Research; 2001.
[22] Stata Statistical Software: Release 10. College Station, TX: StataCorp LP; 2007.
[23] David TR, Packer LE, Heller DC, Chromy JR. 2000 National Household Survey on Drug Abuse: statistical inference. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2002.
[24] Martin LM, Leff M, Calonge N, Garrett C, Nelson DE. Validation of self-reported chronic conditions and health services in a managed care population. Am J Prev Med 2000;18:215–8.
[25] Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 2004;57:1096–103.
[26] Bombard JM, Powell KE, Martin LM, Helmick CG, Wilson WH. Validity and reliability of self-reported arthritis: Georgia Senior Centers, 2000–2001. Am J Prev Med 2005;28:251–8.

Published in final edited form as:
J Hosp Med. 2012 ; 7(6): 470–475. doi:10.1002/jhm.1940.

Predictors of medication adherence post-discharge: the impact of patient age, insurance status and prior adherence

Marya J. Cohen, MD, MPH,
Massachusetts General Hospital, Harvard Medical School

Shimon Shaykevich, MS,
Brigham and Women's Hospital

Courtney Cawthon, MPH,
Vanderbilt Center for Health Services Research

Sunil Kripalani, MD, M.Sc.,
Section of Hospital Medicine, Division of General Internal Medicine and Public Health, Department of Medicine, Vanderbilt University; Vanderbilt Center for Health Services Research

Michael K. Paasche-Orlow, MD, MA, MPH, and
Boston Medical Center, Boston University School of Medicine

Jeffrey L. Schnipper, MD, MPH
BWH Hospitalist Service and Division of General Medicine, Brigham and Women's Hospital, Harvard Medical School

Abstract

Background—Optimizing post-discharge medication adherence is a target for avoiding adverse events. Nevertheless, few studies have focused on predictors of post-discharge medication adherence.

Methods—The Pharmacist Intervention for Low Literacy in Cardiovascular Disease (PILL-CVD) study used counseling and follow-up to improve post-discharge medication safety. In this secondary data analysis, we analyzed predictors of self-reported medication adherence after discharge. Based on an interview at 30 days post-discharge, an adherence score was calculated as the mean adherence in the previous week of all regularly scheduled medications. Multivariable linear regression was used to determine the independent predictors of post-discharge adherence.

Results—The mean age of the 646 included patients was 61.2 years, and they were prescribed an average of 8 daily medications. The mean post-discharge adherence score was 95% (SD = 10.2%). For every 10 year increase in age, there was a 1% absolute increase in post-discharge adherence (95% CI 0.4% −2.0%). Compared to patients with private insurance, patients with Medicaid were 4.5% less adherent (95% CI −7.6% to −1.4%). For every 1-point increase in baseline medication adherence score, as measured by the 4-item Morisky score, there was a 1.6% absolute increase in post-discharge medication adherence (95% CI 0.8% to 2.4%). Surprisingly, health literacy was not an independent predictor of post-discharge adherence.

Corresponding Author: Marya J. Cohen, MD, MPH1,2, General Medicine Unit, Massachusetts General Homspital, 50 Staniford St., 9th Floor, Boston, MA 02114, mjcohen@partners.org, 617-869-7935 phone, 617 -889-8579 fax.

Meeting Presentations:
SGIM New England Regional Meeting, oral presentation, Boston, MA, March 4, 2011
SGIM National Meeting, poster presentation, Phoenix, AZ, May 6, 2011

Conclusions—In patients hospitalized for cardiovascular disease, predictors of lower medication adherence post-discharge included younger age, Medicaid insurance, and baseline non-adherence. These factors can help predict patients who may benefit from further interventions.

Background

In the outpatient setting, medication adherence (defined as percentage of prescribed medication doses taken by a patient during a specific time period) ranges between 40 and 80 percent for chronic conditions. [1] During acute care hospitalization, changes are often made to patients' medication regimens, which can be confusing and contribute to non-adherence, medication errors, and harmful adverse events. [2] Indeed it is estimated that almost half of patients encounter a medication error after discharge, and approximately 12–17% experience an adverse drug event after returning home. [3–6] It is likely that some of these adverse events may be the result of medication non-adherence. [7] Improved patient-provider communication, systems to reconcile pre- and post-hospitalization medications, as well as development of mechanisms to enhance adherence may prevent many of these errors and have become new targets for quality improvement. [4, 8] Although post-discharge medication adherence is a crucial target for avoiding adverse events and re-hospitalization, few studies have focused on understanding its incidence and predictors, in particular patient demographic factors such as age and insurance status. [9–11]

In addition, few studies have looked at general and post-hospital adherence in a population where health literacy is measured, an important area because medication changes during hospitalization may be particularly confusing for patients with low health literacy. [11, 12] Health literacy is defined as "the degree to which an individual has the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions."[13] Prior outpatient research shows that low health literacy is associated with poor patient understanding of the medication regimen and instructions for medication use, which may contribute to post-discharge medication non-adherence.[14, 15] Understanding the factors associated with post-discharge medication adherence could help refine interventions that are oriented toward improving transitions in care, patient safety, and reducing unnecessary re-hospitalization.

We report here on factors associated with post-discharge medication adherence using data from the Pharmacist Intervention for Low Literacy in Cardiovascular Disease (PILL-CVD) study. [16]

Methods

Study and Participants

PILL-CVD was a federally-funded, two-site randomized controlled trial using pharmacist-assisted medication reconciliation, inpatient pharmacist counseling, low-literacy adherence aids, and telephone follow-up that aimed to decrease rates of serious medication errors after hospital discharge. [16] The study targeted patients with cardiovascular disease (hospitalized on cardiology or general medical or geriatric services for acute coronary syndromes [ACS] or acute decompensated heart failure [ADHF]) at two large academic hospitals, Brigham and Women's Hospital (BWH) and Vanderbilt University Hospital (VUH).

Subjects were eligible for enrollment if they met criteria for ACS or ADHF, were likely to be discharged to home as determined by the primary medical team at the time of study enrollment, and took primary responsibility for administering their medications prior to admission (caregivers could be involved in medication management after discharge).

Exclusion criteria included severe visual or hearing impairment, inability to communicate in English or Spanish, active psychiatric illness, dementia, delirium, illness too severe to participate, lack of a home phone number, being in police custody, or participation in another intensive medication adherence program (e.g., due to renal transplant).

Out of 6416 patients originally screened for possible enrollment, 862 were randomly assigned to receive usual care or usual care plus the intervention, and 851 remained in the study.[16] Both the main study and this secondary data analysis were approved by the Institutional Review Boards of each site.

Baseline measures

Following informed consent and study enrollment, a variety of baseline data were collected on study participants from medical records and patient interview, including primary language, demographic information (age, race, insurance status, income, and education level), cognition (through administration of the 0–5 point MiniCog scale) [17] and level of health literacy (through use of the 0–36 point short form of the Test of Functional Health Literacy in Adults, s-TOFHLA scale). [18] Baseline information was also collected on medication use, including number of preadmission medications, measurement of self-reported adherence prior to admission (using the Morisky scale, a validated 0–4 point questionnaire shown to correlate with disease control and indicative of general patterns of adherence), [19] and a medication understanding score, adapted from other instruments, which quantifies understanding of the indication, dose, and frequency of up to 5 randomly selected preadmission medications on a 0–3 point scale. [16,20, 21]

Outcome measures

Outcomes were collected 30 days post-discharge through a structured questionnaire, administered by telephone. Only patients who completed this call are included in the present analysis. Post-discharge medication adherence was assessed by asking patients to report the number of days out of the previous week they had taken each medication from their post-discharge regimen exactly as prescribed.[22] A score was calculated for each medication as the proportion of adherent days (e.g., if a patient reported missing 2 days of a medication in the previous week, then adherence would be 5/7 or 71%). A global post-discharge adherence score was then derived for each patient by averaging the adherence score for all regularly-scheduled medications. This quantitative measure focused on adherence to medications patients knew they should be taking and did not measure medication discrepancies (sometimes termed “unintentional non-adherence”).

Analysis

Patient characteristics were summarized and reported using simple descriptive statistics. Candidate predictors of post-discharge medication adherence were chosen *a priori* from patient characteristics assessed during hospital admission. These included patient age, gender, race, ethnicity, marital status, insurance, years of education, presence of PCP, study site, number of preadmission medications, medication understanding, baseline adherence, cognition, and health literacy. Unadjusted results were calculated using univariable linear regression with each patient’s adherence score as the dependent variable and each predictor as the independent variable. Adjusted results were then derived using multivariable linear regression with all the candidate predictors in the model.

Lastly, because of missing data for some predictors, in particular baseline adherence and medication understanding, multiple imputation techniques were used to impute missing data and increase statistical power. [23] We used the Markov Chain Monte Carlo (MCMC) method for multiple imputation, which generally assumes that the data came from a normal

distribution and that the missing data are missing at random. Because of the essentially normal distribution of the data, and because the amount of missing data was so small (<1% for almost all variables, 5% for baseline adherence, and 8% for medication understanding), we expected little bias and present the complete case analysis, which maximized statistical power.

Two-sided p values < 0.05 were considered significant, and SAS version 9.2 (Cary, NC) was used for all analyses.

Results

Table 1 shows descriptive baseline patient characteristics of study sample (responders) as well as non-responders at 30 days. For the responders, the mean age of the 646 patients was 61.2 years, 94.7% were insured, and 19.3% had inadequate or marginal health literacy. Patients were prescribed an average of 8 preadmission medications. Most patients (92.3%) had a regular primary care physician prior to admission. Non-responders had non-significant trends towards having lower health literacy, medication understanding, and baseline medication adherence.

The average post-discharge adherence score was 95% (SD = 10.2%), and less than 10% of patients had an adherence score of less than 85%; overall the distribution was left-skewed. Table 2 illustrates crude and adjusted parameter estimates for variables in the model. Table 3 shows significant findings in the fully adjusted model, which used multiple imputation techniques to account for missing data.

Intervention arm was of borderline statistical significance in predicting post-discharge adherence (p=0.052), and so was removed from the final model. Study site, age, insurance, and baseline adherence were the only significant independent predictors of post-discharge adherence in the fully adjusted model (Table 3). For example, for every 10 year increase in age, patients had, on average, an adjusted 1% absolute increase in their adherence score (95% CI 0.4% to 2.0%). For every 1-point increase in baseline medication adherence (based on the Morisky scale), there was a 1.6% absolute increase in medication adherence (95% CI 0.8% to 2.4%). In unadjusted analyses, patients with Medicaid were less adherent with medications after discharge than were patients with private insurance. This difference became non-significant in adjusted analyses, but when analyses were repeated using multiple imputation techniques, the results again became statistically significant – Medicaid insurance was associated with a 4.5% absolute decrease in post-discharge adherence compared with private insurance (95% CI −7.6% to −1.4%). Study site (specifically, Brigham and Women's Hospital) was also a significant predictor of greater post-discharge medication adherence. Years of education was a significant predictor of adherence in unadjusted analyses but was not an independent predictor when adjusted for other factors. When baseline adherence was removed from the multiple imputation model, there were no changes in which factors were significant predictors of adherence.

Discussion

In this study, we found that low baseline adherence, younger age, Medicaid insurance, and study site were significant predictors of lower 30-day medication adherence. Of particular interest is our finding regarding baseline adherence, a simple measure to obtain on hospitalized patients. It is notable that in our study, education was not an independent significant predictor of post-discharge adherence, even when baseline adherence was removed from the model. The same is true for medication understanding, cognitive function, and health literacy.

Older patients appeared more adherent with medications in the month after hospital discharge, perhaps reflecting increased interaction with the healthcare system (appointments, number of physician interactions), a greater belief in the importance of chronic medication management, or a higher level of experience with managing medications. A similar relationship between age and adherence has been shown in outpatient studies of patients with hypertension, diabetes, and other chronic diseases. [24–27]

Medicaid patients may be less likely to remain adherent because of the plan's limited coverage of medications relative to patients' ability to pay. For example, Medicaid in Tennessee covers the first 5 generic medications at no cost to the patient but has copayments for additional medications and for brand name drugs. Medicaid in Massachusetts has copayments of $1 to $3 for each medication. Alternatively, Medicaid insurance may be a marker for other patient characteristics associated with low adherence for which we were not fully able to adjust.

Site differences were also notable in this study; these differences could have been due to differences in insurance coverage in Tennessee vs. Massachusetts (which has near-universal coverage), differences in types of insurance (e.g., fewer patients at Brigham and Women's Hospital had Medicaid than at Vanderbilt), cultural and geographic differences between the two locations, or other differences in transitional care between the two sites.

This study corroborates previous literature on medication adherence (specifically unintentional non-adherence) in the outpatient setting, [4, 8–11] for example on the association of younger age with low adherence in certain populations. On the other hand, it may contrast with previous literature which has sometimes shown a relationship between patient education or health literacy and medication adherence. [14, 15, 28–35] However, previous studies have not focused on the transition from inpatient to outpatient settings. Perhaps intensive medication education in the hospital, even under usual care, mitigates the effects of these factors on post-discharge adherence. Finally, baseline adherence seems to correlate with post-discharge adherence, a finding which makes intuitive sense and has been previously reported for specific medications.[36]

There are several limitations to this study. Although large, the study was performed at only two clinical sites where most patients were white and fairly well-educated, perhaps because patients admitted to a tertiary care center with ACS or ADHF are more affluent than general medical inpatients as a whole; this may limit generalizability. Post-discharge medication adherence might have been higher than in other patient populations given the nature of the population, possible loss-to-follow-up bias, and the fact that half of the subjects received an intervention designed to improve medication management after discharge; such low rates of non-adherence in our study may have reduced our ability to detect important predictors in our models. In addition, the period of follow-up was 30 days, thus limiting our findings to short-term post-discharge medication adherence. Post-discharge medication adherence was based on patient self-report, which not only assumed that the patient was still managing his/her own medications after discharge, but may also be susceptible to both recall and social acceptability bias, which might overestimate our adherence scores, again limiting our ability to detect important predictors of non-adherence. However, other studies have shown a good correlation between self-reported medication adherence and other more objective measures [37, 38] and recall was only for 7 days, a measure used previously in the literature[39, 40] and one designed to reduce recall bias. Systematic under-reporting in certain patient populations is less likely but possible.

In the future, research should focus on targeting patients who have low baseline adherence to evaluate the effects of various interventions on post-discharge medication outcomes.

Repeating the study in a population with a high prevalence of low health literacy might be illuminating given that previous studies have shown that patients with low health literacy have less ability to identify their medications and have less refill adherence. [29, 30]

In conclusion, in patients hospitalized with cardiovascular disease, predictors of lower post-discharge adherence include younger age, Medicaid insurance, and low baseline adherence. It may be prudent to assess baseline adherence and insurance type in hospitalized patients in order to identify those who may benefit from additional assistance to improve medication adherence and medication safety during transitions in care.

Acknowledgments

Dr. Schnipper had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med. Aug 4; 2005 353(5):487–497. [PubMed: 16079372]
2. Coleman EA, Smith JD, Raha D, Min SJ. Posthospital medication discrepancies: prevalence and contributing factors. Arch Intern Med. Sep 12; 2005 165(16):1842–1847. [PubMed: 16157827]
3. Cua YM, Kripalani S. Medication use in the transition from hospital to home. Ann Acad Med Singapore. Feb; 2008 37(2):136–136. [PubMed: 18327350]
4. Moore C, Wisnivesky J, Williams S, McGinn T. Medical errors related to discontinuity of care from an inpatient to an outpatient setting. J Gen Intern Med. Aug; 2003 18(8):646–651. [PubMed: 12911647]
5. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. Ann Intern Med. Feb 4; 2003 138(3): 161–167. [PubMed: 12558354]
6. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. Adverse drug events occurring following hospital discharge. J Gen Intern Med. Apr; 2005 20(4):317–323. [PubMed: 15857487]
7. Schnipper JL, Kirwin JL, Cotugno MC, et al. Role of pharmacist counseling in preventing adverse drug events after hospitalization. Arch Intern Med. Mar 13; 2006 166(5):565–571. [PubMed: 16534045]
8. Vira T, Colquhoun M, Etchells E. Reconcilable differences: correcting medication errors at hospital admission and discharge. Qual Saf Health Care. Apr; 2006 15(2):122–126. [PubMed: 16585113]
9. Hassan M, Lage MJ. Risk of rehospitalization among bipolar disorder patients who are nonadherent to antipsychotic therapy after hospital discharge. Am J Health Syst Pharm. Feb 15; 2009 66(4):358–365. [PubMed: 19202045]
10. Mansur N, Weiss A, Hoffman A, Gruenewald T, Beloosesky Y. Continuity and adherence to long-term drug treatment by geriatric patients after hospital discharge: a prospective cohort study. Drugs Aging. 2008; 25(10):861–870. [PubMed: 18808210]
11. Kripalani S, Henderson LE, Jacobson TA, Vaccarino V. Medication use among inner-city patients after hospital discharge: patient-reported barriers and solutions. Mayo Clin Proc. May; 2008 83(5): 529–535. [PubMed: 18452681]
12. Lindquist LA, Go L, Fleisher J, Jain N, Friesema E, Baker DW. Relationship of health literacy to intentional and unintentional non-adherence of hospital discharge medications. J Gen Intern Med. Feb; 27(2):173–178. [PubMed: 21971600]
13. [Accessed February 15, 2012] http://www.healthypeople.gov/Document/pdf/uih/2010uih.pdf
14. Davis TC, Wolf MS, Bass PF 3rd, et al. Literacy and misunderstanding prescription drug labels. Ann Intern Med. Dec 19; 2006 145(12):887–894. [PubMed: 17135578]
15. Kripalani S, Henderson LE, Chiu EY, Robertson R, Kolm P, Jacobson TA. Predictors of medication self-management skill in a low-literacy population. J Gen Intern Med. Aug; 2006 21(8):852–856. [PubMed: 16881946]

16. Schnipper JL, et al. Rationale and design of the Pharmacist Intervention for Low Literacy in Cardiovascular Disease (PILL-CVD) study. Circ Cardiovasc Qual Outcomes. 2010; 3(2):212–219. [PubMed: 20233982]
17. Borson S, Scanlan JM, Watanabe J, Tu SP, Lessig M. Simplifying detection of cognitive impairment: comparison of the Mini-Cog and Mini-Mental State Examination in a multiethnic sample. J Am Geriatr Soc. May; 2005 53(5):871–874. [PubMed: 15877567]
18. Nurss, JR. Short Test of Functional Health Literacy in Adults. Snow Camp, NC: Peppercorn Books and Press; 1998.
19. Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich). May; 2008 10(5):348–354. [PubMed: 18453793]
20. Marvanova MRC, Eden SK, Cawthon C, Schnipper JL, Kripalani S. Health literacy and medication understanding among hospitalized adults. J Hosp Med. In press currently.
21. Marvanova M, Roumie CL, Eden SK, Cawthon C, Schnipper JL, Kripalani S. Health literacy and medication understanding among hospitalized adults. J Hosp Med. Nov.6(9):487. [PubMed: 22139905]
22. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care. Jul; 2000 23(7):943–950. [PubMed: 10895844]
23. Rubin, DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons, Inc; 1987.
24. Hinkin CH, Hardy DJ, Mason KI, et al. Medication adherence in HIV-infected adults: effect of patient age, cognitive status, and substance abuse. AIDS. Jan 1; 2004 18(Suppl 1):S19–25. [PubMed: 15075494]
25. Wong MC, Jiang JY, Griffiths SM. Factors associated with antihypertensive drug compliance in 83,884 Chinese patients: a cohort study. J Epidemiol Community Health. Oct; 64(10):895–901. [PubMed: 20693495]
26. Wong MC, Kong AP, So WY, Jiang JY, Chan JC, Griffiths SM. Adherence to Oral Hypoglycemic Agents in 26 782 Chinese Patients: A Cohort Study. J Clin Pharmacol. Nov 23.
27. Gazmararian J, Jacobson KL, Pan Y, Schmotzer B, Kripalani S. Effect of a pharmacy-based health literacy intervention and patient characteristics on medication refill adherence in an urban health system. Ann Pharmacother. Jan; 44(1):80–87. [PubMed: 20028960]
28. Kalichman SC, Ramachandran B, Catz S. Adherence to combination antiretroviral therapies in HIV patients of low health literacy. J Gen Intern Med. May; 1999 14(5):267–273. [PubMed: 10337035]
29. Gazmararian JA, Kripalani S, Miller MJ, Echt KV, Ren J, Rask K. Factors associated with medication refill adherence in cardiovascular-related diseases: a focus on health literacy. J Gen Intern Med. Dec; 2006 21(12):1215–1221. [PubMed: 17105519]
30. Persell SD, Osborn CY, Richard R, Skripkauskas S, Wolf MS. Limited health literacy is a barrier to medication reconciliation in ambulatory care. J Gen Intern Med. Nov; 2007 22(11):1523–1526. [PubMed: 17786521]
31. Chew LD, Bradley KA, Flum DR, Cornia PB, Koepsell TD. The impact of low health literacy on surgical practice. Am J Surg. Sep; 2004 188(3):250–253. [PubMed: 15450829]
32. Gatti ME, Jacobson KL, Gazmararian JA, Schmotzer B, Kripalani S. Relationships between beliefs about medications and adherence. Am J Health Syst Pharm. Apr 1; 2009 66(7):657–664. [PubMed: 19299373]
33. Fang MC, Machtinger EL, Wang F, Schillinger D. Health literacy and anticoagulation-related outcomes among patients taking warfarin. J Gen Intern Med. Aug; 2006 21(8):841–846. [PubMed: 16881944]
34. Paasche-Orlow MK, Cheng DM, Palepu A, Meli S, Faber V, Samet JH. Health literacy, antiretroviral adherence, and HIV-RNA suppression: a longitudinal perspective. J Gen Intern Med. Aug; 2006 21(8):835–840. [PubMed: 16881943]

35. Platt AB, Localio AR, Brensinger CM, et al. Risk factors for nonadherence to warfarin: results from the IN-RANGE study. Pharmacoepidemiol Drug Saf. Sep; 2008 17(9):853–860. [PubMed: 18271059]

36. Muntner P, Mann DM, Woodward M, et al. Predictors of low clopidogrel adherence following percutaneous coronary intervention. Am J Cardiol. Sep 15; 108(6):822–827. [PubMed: 21741610]

37. Shi L, Liu J, Fonseca V, Walker P, Kalsekar A, Pawaskar M. Correlation between adherence rates measured by MEMS and self-reported questionnaires: a meta-analysis. Health Qual Life Outcomes. 8:99. [PubMed: 20836888]

38. Shi L, Liu J, Koleva Y, Fonseca V, Kalsekar A, Pawaskar M. Concordance of adherence measurement using self-reported adherence questionnaires and medication monitoring devices. Pharmacoeconomics. Dec 1; 28(12):1097–1107. [PubMed: 21080735]

39. Grant RW, Devita NG, Singer DE, Meigs JB. Polypharmacy and medication adherence in patients with type 2 diabetes. Diabetes Care. May; 2003 26(5):1408–1412. [PubMed: 12716797]

40. Grant RW, Devita NG, Singer DE, Meigs JB. Improving adherence and reducing medication discrepancies in patients with diabetes. Ann Pharmacother. Jul-Aug;2003 37(7–8):962–969. [PubMed: 12841801]

Table 1

Baseline characteristics

Characteristic	Total N, 30 day respondents	Value	Total N, nonrespondents	Value
Age, mean in years (SD)	646	61.2 (13.5)	45	55.4 (14.3)
Gender, N (percentage)	646		45	
- Female		272 (42.1)		18 (40.0)
- Male		374 (57.9)		27 (60.0)
Race, N (percentage)	643		45	
- White		511 (79.5)		32 (71.1)
- Black		104 (16.2)		11 (24.4)
- Other		28 (4.4)		2 (4.4)
Ethnicity, N (percentage)	639		45	
- Hispanic		24 (3.8)		1 (2.2)
- Not Hispanic		615 (96.2)		44 (97.8)
Marital status, N (percentage)	646		45	
- Married/cohabitate		382 (59.1)		20 (44.4)
- Separated/divorced		118 (18.3)		11 (24.4)
- Widowed		81 (12.5)		5 (11.1)
- Never married		65 (10.1)		9 (2.0)
Insurance type, N (percentage)	646		45	
- Medicaid		53 (8.2)		5 (11.1)
- Medicare		270 (41.8)		13 (28.9)
- Private		289 (44.7)		19 (42.2)
- Self-pay		34 (5.3)		8 (17.8)
Years of education, mean in years (SD)	643	14.0 (3.1)	45	13.3 (2.7)
Presence of PCP prior to admission, N (percentage)	646		45	
- Yes		596 (92.3)		38 (84.4)
- No		50 (7.74)		7 (15.6)
Site, N (percentage)	646		45	
- Site 1		358 (55.4)		8 (17.8)
- Site 2		288 (44.6)		37 (82.2)
Number of preadmission medications, mean number (SD)	641	7.8 (4.8)	45	7.7 (5.4)
Medication Understanding Score, mean (SD) *	597	2.4 (0.5)	40	2.2 (0.62)
Health Literacy (STOFHLA) score, mean (SD) †	642	29.1 (8.9)	45	26.0 (12.0)
Baseline adherence (SD) ‡	613	2.7 (1.1)	45	2.4 (1.2)

Characteristic	Total N, 30 day respondents	Value	Total N, nonrespondents	Value
Minicog score, N (percentage)//	646		45	
- Demented		63 (9.8)		5 (11.1)
- Not demented		583 (90.2)		40 (88.9)

* 0–3, with 3 indicating better understanding

† 0–36, with higher scores indicating higher health literacy

‡ 0–4, with 4 indicating higher baseline adherence

// 0–5, with higher scores indicating better cognition. A score < 3indicates dementia

Table 2

Crude and Adjusted Measurements

Predictor	Crude parameter estimate (beta) with 95% confidence intervals	P value	Adjusted parameter estimate (beta) with 95% confidence intervals	P value
Age per 10 years	0.010 (0.007–0.020)	<0.0001	0.010 (0.002, 0.020)	0.018
Male gender	0.012 (−0.004, 0.028)	0.137	0.003 (−0.014, 0.020)	0.727
Race/ethnicity				
- White	0.011 (−0.009, 0.031)	0.266	Ref	Ref
- Black	−0.017 (−0.038, 0.005)	0.13	0.006 (−0.017, 0.030)	0.598
- Other	0.010 (−0.029, 0.049)	0.599	0.017 (−0.027, 0.062)	0.446
Hispanic/Latino	0.005 (−0.037, 0.047)	0.803	0.036 (−0.013, 0.085)	0.149
Marital status				
- Married/cohabitate	0.006 (−0.011, 0.022)	0.500	Ref	Ref
- Separated/divorced	−0.005 (−0.025, 0.016)	0.664	0.009 (−0.014, 0.031)	0.446
- Widowed	0.001 (−0.023, 0.025)	0.922	−0.013 (−0.039, 0.013)	0.338
- Never married	−0.009 (−0.035, 0.018)	0.515	−0.004 (−0.033, 0.025)	0.784
Insurance type				
- Private	0.008 (−0.008, 0.024)	0.347	Ref	Ref
- Medicaid	−0.046 (−0.075, −0.018)	0.002	−0.026 (−0.058, 0.007)	0.121
- Medicare	0.012 (−0.004, 0.028)	0.138	−0.002 (−0.023, 0.018)	0.844
- Self-pay	−0.027 (−0.062, 0.008)	0.135	−0.029 (−0.073, 0.015)	0.202
Years of education	0.003 (0.0003, 0.005)	0.028	0.0001 (−0.003, 0.003)	0.949
Presence of PCP prior to admission	0.007 (−0.022, 0.037)	0.630	0.002 (−0.032, 0.036)	0.888
Site	−0.050 (−0.065, −0.034)	<0.0001	−0.038 (−0.056, −0.021)	<0.0001
Number of preadmission medications	−0.0003 (−0.002, 0.001)	0.684	−0.0001 (−0.002, 0.002)	0.918
Med understanding score per point	0.007 (−0.009, 0.023)	0.390	0.006 (−0.011, 0.023)	0.513
Health literacy (STOFHLA) score per 10 points	0.0006 (−0.008, 0.01)	0.897	0.003 (−0.008, 0.01)	0.644
Baseline adherence per point	0.023 (0.016, 0.031)	<0.0001	0.017 (0.009, 0.024)	<0.0001
Cognitive function	0.004 (−0.022, 0.031)	0.757	0.008 (−0.019, 0.036)	0.549

• For crude estimates, value is category versus absence of parameter in bivariate testing

• For adjusted estimates of categorical variables, value is each category compared to referent category

• Beta-coefficient represents absolute change in adherence (e.g., 0.010 for age means a 1% absolute increase in adherence for every 10 year increase in patient age)

Table 3

Significant results in adjusted analyses with multiple imputation

Predictor	Parameter estimate (beta) with 95% confidence intervals	P value
Age per 10 years	0.010 (0.004, 0.020)	0.004
Insurance type		
- Private	Ref	Ref
- Medicaid	−0.045 (−0.076, −0.014)	0.005
- Medicare	−0.010 (−0.030, 0.010)	0.333
- Self-pay	−0.013 (−0.050, 0.025)	0.512
Site	−0.036 (−0.053, −0.019)	<0.0001
Baseline adherence per point	0.016 (0.008, 0.024)	<0.0001

• Total observations 646; 67 with missing values

• All variables adjusted for gender, race, cognitive function, number of preadmission medications, marital status, health literacy score, medication understanding score, presence of PCP, years of school, Hispanic/Latino ethnicity.

A Doctor's Response to Torture

Stepping off the plane in an unfamiliar country, jet-lagged, and unable to speak a word of the native language, I was a little disoriented. I met my patient—whom I will call Rashid—at the hotel in less than an hour of my arrival. He was apprehensive. We spent the evening in a relaxed, informal atmosphere with a group of attorneys. Despite the fact that I was a white, non-Muslim, American woman, we were able to make a connection. Rashid took us to a beautiful beach, peppered with traditional wooden fishing boats, and he gave local children a few coins to pick some baobab fruit for us to sample from a very unusual looking tree. The attorneys, Rashid, and I listened to his music and teased him about his taste in Western hip-hop.

As a physician evaluator for Physicians for Human Rights, I had been asked to interview Rashid. There was no clinical space in which to perform this evaluation, so we used my hotel room. I quickly learned the extent of Rashid's profound suffering. He described himself as a ghost wandering around the town in isolation, unable to eat or sleep. He repeatedly asked me if he was "crazy," because that is how he perceived himself. He detailed the horrors of his arrest, during which he was beaten so badly that he was admitted to a hospital with multiple fractures and internal injuries. He described how he was kidnapped from his hospital bed and survived a 5-year ordeal in U.S. custody in multiple detention sites, including Afghanistan. In the stifling, humid hotel room, he detailed the methods of his torture: severe beatings, prolonged painful stress positions, prolonged solitary confinement, forced nakedness and humiliation, sleep deprivation, withholding of food, sexual assault (anal rape and sodomy), forced intravenous medication during interrogations that he thought might be a "truth serum," and painful shackling. At times, he was denied medical and psychiatric care.

As the hours wore on, Rashid's story became almost unbearable. There were times when we both broke down in tears. Rashid described being forced to lie on a wet mat, naked and handcuffed, having cold water poured on him while being kicked and beaten, and then rolled up in the wet mat "like a corpse"—cold, wet, and shivering. In the "water room," men attempted to insert the spout of a water jug into his anus. He reported that his arms were chained to an overhead pipe while he was in a standing position for what he estimates was 4 days. His heels could not touch the floor and he developed severe back and shoulder pain. He was not allowed to use the toilet, and loud music blared the entire time. ("It is just death.") He went on to describe being locked naked in a "coffin"; he could not move and it was difficult to breathe. At one point during his detention, his flesh started rotting under a cast that was left unattended for too long. His request for medical care was ignored until he protested by going on a hunger strike for 6 days.

Rashid's prominent symptoms included extreme sleep disturbance, sadness, loss of appetite with substantial weight loss, and difficulty interacting with other people, (including family and friends), resulting in profound isolation. He has not been able to work as a result of his debilitating psychological symptoms and has no means to support himself. The lack of self-sufficiency has caused further depression, feelings of inadequacy, and shame and humiliation when he has to rely on his family for basic needs. Rashid told me he wakes up at 2 a.m. and takes walks. ("My head feels empty, like an empty box.") His life has unraveled since his return from U.S. custody, and he is unable to return to his former level of functioning and reintegrate back into his family and community. He meets diagnostic criteria for posttraumatic stress disorder and major depression, but those Western-based diagnoses do not adequately characterize his palpable suffering.

In the evening, Rashid borrowed a car and we decompressed on a drive to the beautiful bush and beaches. I had to abandon the Western doctor–patient boundaries. I learned about local life and customs and tasted the local food. I photographed the old colonial buildings and monkeys and camels. I then met his mother and sister, at his request. We sat on the floor of a simple but welcoming cement structure, were served sweets and coffee, and were treated as warmly as if we were family or close friends. I explained to Rashid's family that he wasn't "crazy" and that there was hope for recovery. Neither Rashid nor his family knew of posttraumatic stress disorder, or understood the link between the torture and his symptoms. Their gratitude for my simple explanation was humbling. We conducted our final interviews on the beach at Rashid's request. After his 5 years in confinement, this vast expanse of beauty was a place of solace and comfort for him. He would stop every so often and illustrate a point by drawing in the sand, such as when he drew a rough reproduction of a cage in which he had been confined. It was difficult to leave Rashid after having uncovered such desperate need and having no available resources to alleviate his suffering. My fears were intensified by the fact that Rashid's suffering was not the aftermath of a natural disaster, but one deliberately inflicted by other humans. After departure, I received a text message from Rashid. "Thank u my Dr. U help me very much and my brain it gonna be OK."

This unique and complicated situation demanded creative and unorthodox measures. The combined collaboration of Physicians for Human Rights and a local organization made initial treatment for this man a reality. After I left, the United Nations Voluntary Fund for Victims of Torture agreed to fund treatment and rehabilitation services, and a psychologist traveled from a neighboring coun-

try to provide mental health treatment. An antidepressant was sent by post, and I initially gave instructions and monitored Rashid's progress by cell phone. I continue to have regular phone contact with him to provide support and to monitor symptoms while trying to organize more regular treatment. I have seen improvement in his self-esteem because he is better able to understand the root cause of his symptoms. I have detected something new in our communication—hope. Rashid is now hopeful about his recovery and future. And I am hopeful that it is possible to repair the wounds my country has inflicted.

Rashid was given a document upon his release that confirmed his confinement as well as his innocence. On the basis of my evaluation, I have a high degree of medical certainty that he was tortured while in U.S. custody. Nonetheless, neither he nor I know the identity of the perpetrators.

Sondra S. Crosby, MD
Boston Medical Center
Boston, MA 02118

Requests for Single Reprints: Sondra S. Crosby, MD, Boston Medical Center, 801 Massachusetts Avenue, Crosstown 2, Boston, MA 02118; e-mail, scrosby@bu.edu.

Ann Intern Med. 2012;156:471-472.

The Impact of a 25-Cent-Per-Drink Alcohol Tax Increase

James I. Daley, MPH, Mandy A. Stahre, MPH, Frank J. Chaloupka, PhD, Timothy S. Naimi, MD, MPH

Background: Excessive alcohol consumption causes 79,000 deaths annually in the U.S., shortening the lives of those who die from it by approximately 30 years. Although alcohol taxation is an effective measure to reduce excessive consumption and related harm, some argue that increasing alcohol taxes places an unfair economic burden on "responsible" drinkers and socially disadvantaged people.

Purpose: To examine the impact of a hypothetic tax increase based on alcohol consumption and sociodemographic characteristics of current drinkers, individually and in aggregate.

Methods: Data from the 2008 Behavioral Risk Factor Surveillance System survey were analyzed from 2010 to 2011 to determine the net financial impact of a hypothetic 25-cent-per-drink tax increase on current drinkers in the U.S. Higher-risk drinkers were defined as those whose past-30-day consumption included binge drinking, heavy drinking, drinking in excess of the U.S. Dietary Guidelines, and alcohol-impaired driving.

Results: Of U.S. adults who consumed alcohol in the past 30 days, 50.4% (or approximately 25% of the total U.S. population) were classified as higher-risk drinkers. The tax increase would result in a 9.2% reduction in alcohol consumption, including an 11.4% reduction in heavy drinking. Compared with lower-risk drinkers, higher-risk drinkers would pay 4.7 times more in net increased annual per capita taxes, and 82.7% of the net increased annual aggregate taxes. Lower-risk drinkers would pay less than $30 in net increased taxes annually. In aggregate, groups who paid the most in net tax increases included those who were white, male, aged 21–50 years, earning ≥$50,000 per year, employed, and had a college degree.

Conclusions: A 25-cent-per-drink alcohol tax increase would reduce excessive drinking, and higher-risk drinkers would pay the substantial majority of the net tax increase.
(Am J Prev Med 2012;42(4):382–389)

Background

Excessive alcohol consumption is a leading cause of preventable death[1] attributed to approximately 79,000 deaths annually in the U.S., shortening the lives of those who die of it by an average of 30 years.[2] Drinking is also a major risk factor for a variety of adverse health outcomes, such as unintentional injuries, violence, unintended pregnancy, and cardiovascular disease.[3–7] Alcohol is also a major cause of social problems (e.g., child neglect, divorce); legal and criminal justice problems; and economic costs.[4]

Increasing the price of alcohol through alcohol excise taxes is an effective means of reducing excessive drinking, and it is considered the most important public health intervention to reduce alcohol-related harm.[8–10] A meta-analysis of 50 publications found that doubling the alcohol excise tax would reduce alcohol-related mortality by an average of 35%, traffic crash deaths by 11%, sexually transmitted disease by 6%, violence by 2%, and crime by 1.4%.[10] Further, a comprehensive review found an inverse relationship between alcohol price and consumption, and determined that a 10% increase in alcohol prices would result in a 3%–10% decrease in overall consumption.[9]

Despite strong evidence of public health benefit, there have been few recent alcohol tax increases, and many initiatives to raise them have been defeated on the basis of arguments about the economic fairness of alcohol taxes. One argument against raising alcohol taxes is that such

From the Boston University School of Medicine (Daley); Section of General Internal Medicine (Naimi), Boston Medical Center, Boston, Massachusetts; Alcohol Epidemiology Program (Stahre), University of Minnesota School of Public Health, Minneapolis, Minnesota; and Health Policy Center (Chaloupka), University of Illinois at Chicago, Chicago, Illinois

Address correspondence to: Timothy S. Naimi, PhD, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Ave., Rm. 2079, Boston MA 02118. E-mail: tim.naimi@bmc.org.

0749-3797/$36.00

doi: 10.1016/j.amepre.2011.12.008

increases will place a large and unfair financial burden on "responsible" (i.e., lower-risk) drinkers. Also, there is concern about how much those with lower socioeconomic status (i.e., racial and ethnic minorities, those with lower incomes, less education) would have to pay compared with other groups. To explore these issues further, data from the Behavioral Risk Factor Surveillance System (BRFSS) survey were analyzed to examine the effect of a hypothetic 25-cent-per-drink tax increase on U.S. adults who drink alcohol. The purpose of the present study was to assess the net financial impact, both individually and in aggregate, of a hypothetic tax increase on the basis of alcohol consumption characteristics and socioeconomic factors among current drinkers in the U.S.

Methods

Survey

Data for the current study came from the 2008 BRFSS. A detailed overview of the BRFSS including survey methods and information on data weighting is available at www.cdc.gov/brfss/ and www.cdc.gov/brfss/technical_infodata/quality.htm. In brief, the BRFSS is a state-based random-digit-dial telephone survey of people aged ≥18 years that is overseen by the CDC and conducted monthly in all states, the District of Columbia, and some territories. Analysis was limited to the 50 U.S. states. The 2008 BRFSS had a response rate of 53.3%. A total of 411,736 respondents, including 200,587 reporting alcohol consumption in the previous 30 days, constituted the study population. BRFSS data were weighted on the basis of age, gender, and race/ethnicity to be representative of the U.S. population aged ≥18 years.

Measures

The BRFSS "core" (i.e., questions asked of all drinkers in all states) alcohol questions were used to assess drinking frequency and the average number of drinks consumed during drinking days during the past 30 days among current drinking U.S. adults aged ≥18 years. The core BRFSS also collects information on the frequency of binge drinking (i.e., consuming five or more drinks for men or four or more drinks for women per occasion in the past 30 days),[11] the greatest number of drinks consumed in the past 30 days, and alcohol-impaired driving (i.e., the number of times a person drove after having "perhaps too much to drink").

To determine average daily alcohol consumption by an individual, a participant's frequency of drinking days in the past 30 days was multiplied by their usual number of drinks consumed per drinking day and divided by 30. For binge drinkers, their average daily alcohol consumption was adjusted to account for self-reported estimates of binge drinking, using the indexing method originally described by Armor and Polich[12] and applied to the BRFSS data.[13] For the current study, men reporting a maximum drink value of five or more and women reporting a maximum drink value of four or more were classified as binge drinkers.[11] Heavy drinking was defined as an average daily consumption of more than two drinks for men or an average daily consumption of more than one drink for women.[2]

"Higher-risk" drinkers were defined as those with any of the following: drinking above U.S. Dietary Guidelines (defined as consumption during drinking days of three or more drinks for men or two or more drinks for women)[14]; binge drinking; alcohol-impaired driving; or heavy drinking. "Lower-risk" drinkers were defined as those not meeting any of the four higher-risk criteria defined above.

Increasing the price of alcohol through a tax would decrease its consumption based on the price elasticity of demand for alcohol. The price elasticity of demand (−0.51) was obtained from a recent meta-analysis.[15] To determine the average price of a standard drink in the U.S., a weighted average of on- and off-premises prices for beer, wine, and spirits was calculated[16] based on beverage-specific per capita consumption in the U.S. in 2007 (the most recent year for which data were available: 52.1% beer, 31.5% liquor, 16.4% wine[17]) and assuming that 75% of alcohol is purchased from off-premises establishments.[18,19] Assuming a 100% pass-through of the tax to the price paid by consumers, 25 cents constituted 18.1% of the average weighted price per drink ($1.38), yielding a 9.2% reduction in consumption. After factoring in the impact of the price increase on average daily consumption, annual alcohol consumption was calculated on a per capita and aggregate basis by multiplying the average number of drinks per day by 365.

The net tax increase was calculated by first determining the current amount of annual federal alcohol taxes already being paid by drinkers (10 cents per drink[7,20]). Total annual taxes paid after a hypothetic 25-cent-per-drink increase were then calculated by multiplying an individual's price elasticity–adjusted annual consumption by 35 cents per drink (10 cents per drink in current taxes plus a hypothetic 25-cent-per-drink increase). The net increase in expenditures was then calculated by subtracting current annual tax expenditures from the total annual tax expenditures after the tax increase (i.e., after adjusting consumption to account for the impact of the higher price). Per capita and aggregate net tax increase–associated revenues were then assessed among a variety of subgroups based on their alcohol consumption characteristics and sociodemographic factors such as age, gender, race/ethnicity, education, income, and employment status. Analyses were performed from May 2010 through August 2011, using SAS, version 9.2.

Results

After weighting, 50.5% of U.S. adults who reported current drinking in the past 30 days constituted the study population (Appendix A, available online at www.ajpmonline.org). Among current drinkers, 50.4% were classified as higher-risk drinkers, and 49.6% were classified as lower-risk drinkers. Among higher-risk drinkers, 58.8% usually consumed three or more drinks per drinking day; 63.0% reported binge drinking, including 25.3% with three or more binge episodes; and 7.5% reported alcohol-impaired driving. A majority (75.2%) of lower-risk drinkers reported usually consuming one drink per drinking day and 80.2% of lower-risk drinkers reported between 1 and 10 drinking days in the past 30 days. Men drank more than women for all consumption measures. Among male higher-risk drinkers, 78.3% usually consumed three or more drinks per drinking day, 20.2% drank a maximum of ten or more drinks on one occasion in the previous 30 days, 77.5%

reported binge drinking, and 10.4% reported impaired driving.

Higher-risk drinkers were further characterized on the basis of the risk criteria that were used to define them as such (Appendix B, available online at www.ajpmonline.org). The most common risk factor reported was exceeding the U.S. Dietary Guidelines (82.9%), followed by binge drinking (63.0%); heavy drinking (26.4%); and impaired driving (7.5%). Overall, higher-risk drinkers reported a mean of 1.8 of the four risk criteria. The mean number of fulfilled risk criteria based on strata of socioeconomic factors including race/ethnicity, education, income, and employment status was similar overall, ranging from 1.7 to 1.9.

As a result of the tax increase, current drinkers would reduce their annual consumption from 319.1 to 289.7 drinks per year (absolute reduction 29.4 drinks annually, relative reduction 9.2%, data not shown). Compared with lower-risk drinkers, higher-risk drinkers would have larger absolute reductions in the number of drinks consumed (48.4 fewer drinks vs 10.2 fewer drinks). Among heavy drinkers (the subset of higher-risk drinkers defined on the basis of average consumption), the tax increase would reduce annual consumption by 58.6 drinks annually, and result in an 11.4% relative reduction in heavy drinking (from 14.0% to 12.4%, or 1.8 million people).

Table 1 provides per capita data regarding total drink consumption and the net additional cost of a hypothetic alcohol tax increase of 25 cents per drink, analyzed by various socioeconomic characteristics. On a per capita basis, higher-risk drinkers would pay 4.7 times as much in net additional taxes compared with lower-risk drinkers. Those exceeding four risk criteria would pay approximately 10 times more than those exceeding only one risk criteria, and approximately 16 times more than those exceeding none (i.e., lower-risk drinkers). Among various strata of lower-risk drinkers, none would pay more than $35.04 per year in net increased taxes. Higher-risk drinkers who would pay the most in additional taxes include men, those with less education, those of lower income, and retired people. Lower-risk drinkers who would pay the most in additional taxes include men, those of white race, college graduates, and those with incomes exceeding $50,000.

Table 2 presents similar data but on an aggregate, rather than per capita, basis. Additional net tax revenue was approximately $7.9 billion annually, and 82.7% ($6.5 billion) of this was to be paid by higher-risk drinkers. Among all drinkers and among higher-risk drinkers, those aged 21–50 years, men, those of white race, college graduates, those earning $50,000 or more per year, and employed people paid the most in terms of aggregate tax increases. Those exceeding three alcohol-related risk criteria (11.5% of all drinkers) paid approximately 48% of tax revenues.

Discussion

Raising the price of alcohol through taxation is a highly effective strategy by which to reduce excessive drinking and related harm. A 25-cent-per-drink tax increase would result in more than a 10% reduction in heavy drinking, which would yield a substantial public health benefit for a behavior that currently leads to approximately 79,000 deaths annually in the U.S.[2] However, it is likely that the morbidity and mortality benefits would exceed those suggested by the change in consumption, because absolute consumption would decrease most among those who drink the most and who incur most alcohol-attributable consequences.[10,21]

In addition, because the risk of alcohol-attributable harm increases exponentially with increasing levels of consumption,[22] even moderate reductions in consumption among those drinking the most leads to relatively large reductions in alcohol-attributable harm. Finally, youth drinkers may experience greater-than-average reductions in alcohol consumption from a tax increase because they tend to be high-risk drinkers and because their price elasticity of demand is typically high. This effect was not modeled, because of a lack of an available meta-analysis of the price elasticity of demand for this group.[10,21]

Higher-risk drinkers (including those who reported binge drinking, consumption during drinking days that exceeded U.S. Dietary Guidelines limits, heavy drinking, or alcohol-impaired driving) would pay approximately five times as much individually and in aggregate compared with lower-risk drinkers, whereas the typical lower-risk drinker would pay less than $30 in net additional taxes annually. This result demonstrates that even though alcohol taxes are applied to all drinks (i.e., alcohol taxes cannot selectively be applied based on who is purchasing the alcohol), they cost higher-risk drinkers considerably more than lower-risk drinkers because of how skewed alcohol consumption is in the U.S.: A relatively small percentage of drinkers consume most of the alcohol, and most lower-risk drinkers do not drink regularly or consume much alcohol. Having higher-risk drinkers pay far more than lower-risk drinkers not only is desirable from a fairness perspective but results in larger absolute reductions in consumption among higher-risk drinkers, which is desirable from a public health perspective.

There is also interest about who pays the most in alcohol taxes based on socioeconomic factors, particularly among lower-risk drinkers who are unlikely to be detrimental to public health and safety. Among lower-risk

Table 1. Drinking behaviors and tax expenditures derived from a 25-cent-per-drink alcohol tax increase

	Higher-risk drinkers[a]			Lower-risk drinkers		
	Average number of drinks per month	Average number of drinks per year	Per capita net annual tax increase ($)	Average number of drinks per month	Average number of drinks per year	Per capita net annual tax increase ($)
Overall	39.6	475.5	114.03	8.4	100.9	24.20
Age (years)						
18–20	38.7	464.9	111.50	3.0	35.9	8.61
21–30	43.9	527.3	126.47	5.0	60.5	14.51
31–40	34.0	408.5	97.98	6.1	73.7	17.68
41–50	37.1	445.7	106.90	7.6	91.5	21.95
51–60	40.1	480.7	115.29	9.2	110.2	26.43
≥61	46.4	556.2	133.39	12.0	144.1	34.56
Gender						
Male	53.9	646.7	155.10	10.7	128.3	30.78
Female	22.6	271.3	65.07	5.4	64.6	15.49
Race/ethnicity						
White/non-Hispanic	38.8	465.2	111.57	9.1	108.9	26.12
Black/non-Hispanic	36.4	438.3	105.12	5.5	66.5	15.95
Other/non-Hispanic	55.9	670.8	160.88	7.0	84.0	20.15
Hispanic	38.3	459.3	110.16	5.6	66.8	16.02
Education						
Some high school	56.7	680.7	163.26	6.3	75.6	18.13
High school graduate	44.0	527.7	126.56	7.1	84.6	20.29
Some college	36.6	439.7	105.46	7.7	92.1	22.09
College graduate	34.3	411.6	98.71	9.6	114.7	27.51
Income ($)						
<25,000	43.7	524.1	125.70	6.8	81.4	19.52
25,000–34,999	39.6	475.7	114.09	7.3	87.3	20.94
35,000–49,999	37.7	452.1	108.43	7.9	94.3	22.62
>50,000	38.7	463.9	111.26	9.2	109.9	26.36
Employment						
Employed	39.0	468.0	112.24	7.9	95.1	22.81
Unemployed	41.0	491.7	117.93	6.0	71.8	17.22
Student	34.1	409.5	98.22	4.4	53.3	12.79
Retired	46.4	556.4	133.44	12.2	146.1	35.04
Number of risk criteria exceeded[b]						
0	—	—	—	8.4	100.9	24.20
1	13.2	158.0	37.89	—	—	—
2	32.8	393.6	94.40	—	—	—

(continued on next page)

Table 1. Drinking behaviors and tax expenditures derived from a 25-cent-per-drink alcohol tax increase *(continued)*

	Higher-risk drinkers[a]			Lower-risk drinkers		
	Average number of drinks per month	Average number of drinks per year	Per capita net annual tax increase ($)	Average number of drinks per month	Average number of drinks per year	Per capita net annual tax increase ($)
3	94.9	1138.4	273.02	—	—	—
4	137.5	1650.2	395.77	—	—	—

Note: Data are for lower- and higher-risk U.S. adult drinkers, by selected characteristics, BRFSS survey, 2008. Data have been modeled to include a decrease in consumption as a result of the increased price of a drink due to a hypothetic 25-cent-per-drink tax, based on the price elasticity of alcohol. BRFSS data were weighted to be representative for U.S. adults aged ≥18 years on the basis of age, gender, and race/ethnicity.

[a]Higher-risk drinkers were defined as anyone reporting one or more of the following risk factors: binge drinking, heavy drinking, drinking in excess of Dietary Guidelines limits, or alcohol-impaired driving. Lower-risk drinkers were defined as not reporting any of the preceding risk factors. Binge drinking was defined as one or more occasions of consuming four or more drinks for women or five or more drinks for men. Heavy drinking was defined as consuming an average of more than two drinks per day for men or more than one drink per day for women. Drinking above Dietary Guidelines was defined as consumption of three or more drinks for men and two or more drinks for women during drinking days. Impaired driving was defined as nonzero response to the question *During the past 30 days, how many times have you driven when you've had perhaps too much to drink?*

[b]The number of the risk criteria exceeded refers to the sum of any of the four alcohol-related risk criteria used to define higher-risk drinking.

BRFSS, Behavioral Risk Factor Surveillance System

drinkers, both in aggregate and on a per capita basis, groups who paid the most for an alcohol tax increase were male, white, relatively affluent and educated, and employed. Therefore, assuming that alcohol tax revenues were used for offsets across the board to other tax obligations, lower-risk drinkers from relatively disadvantaged socioeconomic groups would realize a net economic gain from an alcohol tax increase.

Among higher-risk drinkers, however, those who were relatively poor, less educated, and not employed paid more in per capita tax increases than other groups. Also, it should be noted that any tax increase on those with less income will take a larger proportion of their income than would the same tax on someone earning more, and could thus be potentially regressive in nature. In this case, however, that larger financial impact might lead to greater reductions in drinking and a larger public health benefit for those same individuals.

It is important to acknowledge that the distribution of taxes paid by higher-risk and lower-risk drinkers is dependent on the classification criteria for the two groups. Of the 57.4 million higher-risk drinkers, 17.8 million were considered higher risk solely because their consumption exceeded the U.S. Dietary Guidelines but exceeded no other risk thresholds. This subset of higher-risk drinkers consumed less alcohol overall (and hence paid less in taxes) than other higher-risk drinkers.

Their net per capita taxes paid would increase by $31.30 annually, and in aggregate they would pay an additional $556 million, or 8.5% of the amount that would be paid by the higher-risk group (data not shown). Had they been classified as lower-risk drinkers instead, the remaining higher-risk drinkers would have paid 6.1 times more than lower-risk drinkers in per capita taxes, and accounted for 75% of total tax paid (compared with paying 4.7 times more than lower-risk drinkers and 83% of all additional taxes paid when only those drinking above U.S. Dietary Guidelines are included in the higher-risk group).

Imposing alternative assumptions would have changed the current findings. Had a different tax increase been used, consumption and revenue increases would have changed accordingly (e.g., higher taxes would cause larger decreases in consumption but higher revenue). The overall price elasticity of demand for the current study was based on a recent meta-analysis[15]; however, had meta-analyses of price elasticities for subgroups (e.g., based on income, age, drinking quantity) been available, the observed distribution of tax revenues would have been somewhat different.

For example, assuming a larger price elasticity of demand for those of younger ages or lower income would have resulted in relatively larger reductions in their drinking and a reduced tax burden for them. Although a full pass-through of the tax was modeled (i.e., assuming that 100% of the tax was passed on as a price increase), modeling either a smaller or larger relative pass-through would have resulted in a smaller or larger change in price, which would have affected consumption and tax impacts accordingly.

Characteristics of surveys generally, including BRFSS, also make it likely that the financial impact of the tax increase would differ from our estimates.[23] Specifically, BRFSS estimates of consumption account for less than one third of consumption based on sales tax data,[24,25] suggesting that actual additional revenue from a 25-cent-

Table 2. Net aggregate annual expenditures from a 25-cent-per-drink alcohol tax increase

	Higher-risk drinkers[a]		Lower-risk drinkers		All drinkers	
	Total population, millions	Net annual tax, $ millions	Total population, millions	Net annual tax, $ millions	Total population, millions	Net aggregate annual tax, $ millions
Overall	57.4	6530.9	56.5	1368.0	113.9	7898.9
Age (years)						
18–20	3.0	331.3	1.1	9.6	4.1	340.8
21–30	15.0	1901.4	6.7	97.4	21.7	1998.8
31–40	13.7	1341.3	10.5	185.5	24.2	1526.9
41–50	12.4	1330.3	12.1	265.0	24.5	1595.3
51–60	8.1	932.8	11.5	303.1	19.6	1236.0
≥61	5.2	693.8	14.7	507.3	19.9	1201.1
Gender						
Male	31.0	4813.6	32.2	990.9	63.2	5804.5
Female	26.4	1717.4	24.3	377.1	50.7	2094.5
Race/ethnicity						
White/non-Hispanic	41.4	4621.0	44.0	1148.9	85.4	5769.9
Black/non-Hispanic	4.5	474.9	4.2	66.7	8.7	541.5
Other/non-Hispanic	3.2	511.5	3.3	66.9	6.5	578.4
Hispanic	8.1	886.8	4.7	75.3	12.8	962.1
Education						
Some high school	4.9	805.7	2.4	43.2	7.3	848.9
High school graduate	15.5	1965.5	11.9	241.8	27.4	2207.3
Some college	16.7	1755.8	14.5	321.4	31.2	2077.1
College graduate	20.3	2001.0	27.7	760.7	48.0	2761.7
Income ($)						
<25,000	10.0	1260.9	6.6	128.5	16.6	1389.5
25,000–34,999	5.2	594.6	4.5	94.2	9.7	688.8
35,000–49,999	7.7	836.2	6.9	156.4	14.6	992.6
≥50,000	30.1	3351.7	33.3	878.2	63.4	4229.9
Employment status						
Employed	42.0	4712.0	37.3	850.5	79.3	5562.5
Unemployed	8.3	980.1	6.9	119.2	15.2	1099.4
Student	3.0	298.3	1.5	19.3	4.5	317.5
Retired	4.0	529.2	10.8	376.9	14.8	906.1
Number of risk criteria exceeded[b]						
0	—	—	56.5	1368.0	56.5	1368.0
1	25.8	980.0	—	—	25.8	980.0
2	18.4	1736.6	—	—	18.4	1736.6

(continued on next page)

Table 2. Net aggregate annual expenditures from a 25-cent-per-drink alcohol tax increase *(continued)*

	Higher-risk drinkers[a]		Lower-risk drinkers		All drinkers	
	Total population, millions	Net annual tax, $ millions	Total population, millions	Net annual tax, $ millions	Total population, millions	Net aggregate annual tax, $ millions
3	11.4	3123.6	—	—	11.4	3123.6
4	1.7	690.8	—	—	1.7	690.8

Note: Data are for higher- and lower-risk U.S. adult drinkers, by selected characteristics, BRFSS survey, 2008. Data have been modeled to include a decrease in consumption as a result of the increased price of a drink due to a hypothetic 25-cent-per-drink tax, based on the price elasticity of alcohol. BRFSS data were weighted to be representative for U.S. adults aged ≥18 years on the basis of age, gender, and race/ethnicity.

[a]Higher-risk drinkers were defined as anyone reporting one or more of the following risk factors: binge drinking, heavy drinking, drinking in excess of Dietary Guidelines limits, or alcohol-impaired driving. Lower-risk drinkers were defined as not reporting any of the preceding risk factors. Binge drinking was defined as one or more occasions of consuming four or more drinks for women or five or more drinks for men. Heavy drinking was defined as consuming an average of more than two drinks per day for men or more than one drink per day for women. Drinking above Dietary Guidelines was defined as consumption of three or more drinks for men and two or more drinks for women during drinking days. Impaired driving was defined as a nonzero response to the question *During the past 30 days, how many times have you driven when you've had perhaps too much to drink?*

[b]The number of risk criteria exceeded refers to the sum of any of the four alcohol-related risk criteria used to define higher-risk drinking.

BRFSS, Behavioral Risk Factor Surveillance System

per-drink tax would be approximately $29 billion. In addition, survey respondents are less likely to include those who consume high amounts of alcohol compared with those who consume lower amounts.[26–29] This implies that in the current study, the proportion of alcohol consumption and tax expenditures accounted for by higher-risk drinkers are underestimates relative to consumption and expenditures by lower-risk drinkers.

Despite the lack of any randomized study of low-dose alcohol and any morbidity or mortality outcome, and despite the fact that observational studies of established moderate drinkers are limited by confounding and selection bias, much attention has been paid to possible health benefits of low-level (sometimes termed "moderate") alcohol consumption.[30–33] However, to the extent that there are health benefits associated with moderate drinking, reducing higher-risk alcohol consumption to more moderate levels by means of a tax increase would have the salutary side effect of increasing the number of people to whom such benefit might accrue. Moreover, the reductions in higher-risk drinking following a tax increase would lead to sharp reductions in overall alcohol-related mortality.

A strong theoretic justification for taxes generally, and for alcohol excise taxes in particular, is to recoup the social and economic costs incurred from the sale of alcohol that are not borne by its producers, sellers, or consumers (i.e., the "external" costs of alcohol consumption).[22] These external costs are considerable, and include health-related effects; social problems (e.g., child abuse and neglect, marital problems, alcohol-related crimes); and economic costs (e.g., healthcare costs, legal and criminal justice system costs, lost economic productivity, and higher car insurance premiums).[2,34–37] Currently, the net external costs (external costs minus tax revenues) for alcohol in the U.S. are approximately $1 per drink; remedying this disparity would require raising alcohol taxes by an order of magnitude from present levels.[38–40]

The gap between the societal costs of alcohol sales and its corresponding excise tax rates continues to widen.[41] This is because most alcohol excise taxes at the federal and state levels are based on volume and have not been increased regularly or in large enough increments over time,[42] resulting in an erosion of the value of these taxes in inflation-adjusted terms.[42,43] For example, the federal beer tax has remained unchanged since it was last raised in 1991, and its inflation-adjusted value has eroded by approximately 40% in the interim.[43,44]

Raising alcohol taxes provides multiple benefits in terms of public health and economic fairness.[44] It would reduce harm from the third-leading preventable cause of death in the U.S.,[1] reduce costs to responsible drinkers by lowering tax obligations in other areas, begin to recoup alcohol-related costs to society associated with alcohol sales by moving taxes more closely in line with historical standards, and provide badly needed revenue to cash-strapped state and federal governments at a time when there is extreme reluctance to increase income or property taxes.

This work was supported in part by grant RO1 AA018377 from the National Institutes of Alcohol Abuse and Alcoholism (NIAAA). The content of this manuscript does not necessarily represent the views of NIAAA or the NIH.

No financial disclosures were reported by the authors of this paper.

References

1. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the U.S., 2000. JAMA 2004;291(10):1238–45.
2. CDC. Alcohol-attributable deaths and years of potential life lost—U.S., 2001. Mortal Morb Wkly Rep 2001;53(37);866–70.
3. Brewer RD, Swahn MH. Binge drinking and violence. JAMA 2005;294(5):616–8.
4. Chikritzhs TN, Jonas HA, Stockwell TR, Heale PF, Dietze PM. Mortality and life-years lost due to alcohol: a comparison of acute and chronic causes. Med J Aust 2001;174(6):281–4.
5. Elder R. Involvement by young drivers in fatal alcohol-related motor-vehicle crashes—U.S., 1982–2001. CDC Morb Mortal Wkly Rep 2002;51:1089–91.
6. Naimi TS, Lipscomb LE, Brewer RD, Gilbert BC. Binge drinking in the preconception period and the risk of unintended pregnancy: implications for women and their children. Pediatrics 2003;111(5 Pt 2): 1136–41.
7. DHHS. 10th special report to the U.S. Congress on alcohol and health. Washington DC: DHHS, 2000.
8. Babor T. Alcohol: no ordinary commodity: research and public policy. Oxford: Oxford University Press, 2003.
9. Elder R, Lawrence B, Ferguson A, et al.; Community Preventive Services Task Force. The effectiveness of tax policy interventions for reducing excessive alcohol consumption and related harms. Am J Prev Med 2010;38(2):217–29.
10. Wagenaar A, Tobler A, Komro K. Effects of alcohol tax and price policies on morbidity and mortality: a systematic review. Am J Public Health 2010;100(11):2270–8.
11. National Institute on Alcohol Abuse and Alcoholism. Binge drinking definition. Bethesda: NIAAA Newsletter 3, 2004.
12. Armor DJ, Polich JM. Measurement of alcohol consumption. New York: Gardner Press, 1982.
13. Stahre M, Naimi TS, Brewer RD, Holt J. Measuring average alcohol consumption: the impact of including binge drinks in quantity-frequency calculations. Addiction 2006;101(12):1711–8.
14. U.S. Department of Agriculture. Dietary guidelines for Americans, 2005. Washington DC: U.S. Department of Agriculture, 2005.
15. Wagenaar A, Salois M, Komro K. Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. Addiction 2009;104(2):179–90.
16. IMPACT databank review and forecast: the U.S. spirits, wine, and beer markets (three volume set). New York: M. Shanken Communications, 2009.
17. National Institute on Alcohol Abuse and Alcoholism. Apparent per capita ethanol consumption for the U.S., 1850–2007. NIAA. www.niaaa.nih.gov/Resources/DatabaseResources/QuickFacts/AlcoholSales/Pages/consum01.aspx.
18. Goldammer T. The brewer's handbook: the complete book to brewing beer. 2nd ed. Clifton NJ: Apex Publishers, 2008.
19. Distilled Spirits Council of the U.S. Press release: distilled spirits review: industry proves recession resilient in 2009. www.discus.org/media/press/article.asp?NEWS_ID=569.
20. U.S. Department of the Treasury, Alcohol and Tobacco Tax and Trade Bureau. Alcohol tax and fee rates. www.ttb.gov/tax_audit/atftaxes.shtml.
21. Grossman MC, Arluck GM. Price sensitivity of alcoholic beverages in the U.S.: youth alcohol consumption. Greenwich CT: JAI Press, 1987.
22. Cook PJ. Paying the tab: the economics of alcohol policy. Princeton NJ: Princeton University Press, 2007.
23. Nelson DE, Holtzman DE, Bolen JT, Stanwyck CA, Mack KA. Reliability and validity of measures from the Behavioral Risk Factor Surveillance System (BRFSS). Soz Praventivmed 2001;46(S1): S3–S42.
24. Nelson DE, Naimi TS, Brewer RD, Roeber J. U.S. state alcohol sales compared to survey data, 1993-2006. Addiction 2010;105(9):1589–96.
25. Smith PF, Remington PL, Williamson DF, Anda RF. A comparison of alcohol sales data with survey data on self-reported alcohol use in 21 states. Am J Public Health 1990;80(3):309–12.
26. Dawson DA. Methodological issues in measuring alcohol use. Alcohol Res Health 2003;27(1):18–29.
27. Jousilahti PH, Salomaa VK, Kuulasmaa KN, Niemelä M, Vartiainen E. Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. J Epidemiol Community Health 2005;59(4):310–5.
28. Midanik LH. The validity of self-reported alcohol consumption and alcohol problems: a literature review. Br J Addict 1982;77(4):357–82.
29. Rosengren AD, Wilhelmsen LG, Berglund GD, Elmfeldt D. Non-participants in a general population study of men, with special reference to social and alcoholic problems. Acta Med Scand 1987; 221(3):243–51.
30. Naimi TS, Brown DW, Brewer RD, et al. Cardiovascular risk factors and confounders among nondrinking and moderate-drinking U.S. adults. Am J Prev Med 2005;28(4):369–73.
31. Fillmore KM, Stockwell TK, Chikrizhs T, Bostrom A, Kerr W. Moderate alcohol use and reduced mortality risk: systematic error in prospective studies and new hypotheses. Ann Epidemiol 2007;17(5S):S16–S23.
32. Hansel B, Thomas F, Pannier B, et al. Relationship between alcohol intake, health and social status and cardiovascular risk factors in the urban Paris-Ile-De-France Cohort: is the cardioprotective action of alcohol a myth? Eur J Clin Nutr 2010;64(6):561–8.
33. Wannamethee G, Shaper A. Men who do not drink: a report from the British Regional Heart Study. Int J Epidemiol 1988;17(2):307–16.
34. Navarro HJ, Doran CM, Shakeshaft AP. Measuring costs of alcohol harm to others: a review of the literature. Drug Alcohol Depend 2011;114(2-3):87–99.
35. Bouchery EE, Harwood HJ, Sacks J, Simon CJ, Brewer RD. Economic costs of excessive alcohol consumption in the U.S., 2006. Am J Prev Med 2011;41(5):516–24.
36. Harwood HJ, Fountain D, Livermore G. Economic costs of alcohol abuse and alcoholism. Recent Dev Alcohol 1998;14:307–30.
37. Manning W, Keller E, Newhouse J, et al. The taxes of sin. Do smokers and drinkers pay their way? JAMA 1989;261(11):1604–9.
38. Cook PJ, Moore MJ. The economics of alcohol abuse and alcohol-control policies. Health Aff (Millwood) 2002;21(2):120–33.
39. The Marin Institute. Increasing alcohol taxes—myth v. reality. San Rafael: Marin Institute, Alcohol Industry Watchdog, 2009.
40. Naimi T. The cost of alcohol and its corresponding taxes in the U.S.: a massive public subsidy for excessive drinking and alcohol industries. Am J Prev Med 2011;41(5):546–7.
41. Alcohol Epidemiology Program: Wagenaar AC. Alcohol policies in the U.S.: highlights from the 50 States. Minneapolis: University of Minnesota, 2009.
42. The International Center for Alcohol Policies. Alcohol taxation, in ICAP Reports. Washington DC: ICAP, 2006.
43. Cook PJ. A free lunch. J Drug Policy Anal 2008;1(1):34–9.
44. Chaloupka FJ, Grossman M, Saffer H. The effects of price on alcohol consumption and alcohol-related problems. Alcohol Res Health 2002; 26(1):22–34.

Appendix

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.amepre.2011.12.008.

REPRODUCTIVE HEALTH

RESEARCH Open Access

Disparities in safe sex counseling & behavior among individuals with substance dependence: a cross-sectional study

Meredith M D'Amore[1*], Debbie M Cheng[2,3], Donald Allensworth-Davies[4], Jeffrey H Samet[2,5] and Richard Saitz[2,6]

Abstract

Background: Despite the vast literature examining disparities in medical care, little is known about racial/ethnic and mental health disparities in sexual health care. The objective of this study was to assess disparities in safe sex counseling and resultant behavior among a patient population at risk of negative sexual health outcomes.

Methods: We conducted a cross-sectional analysis among a sample of substance dependent men and women in a metropolitan area in the United States. Multiple logistic regression models were used to explore the relationship between race/ethnicity (non-Hispanic black; Hispanic; non-Hispanic white) and three indicators of mental illness (moderately severe to severe depression; any manic episodes; ≥3 psychotic symptoms) with two self-reported outcomes: receipt of safe sex counseling from a primary care physician and having practiced safer sex because of counseling.

Results: Among 275 substance-dependent adults, approximately 71% (195/275) reported ever being counseled by their regular doctor about safe sex. Among these 195 subjects, 76% (149/195) reported practicing safer sex because of this advice. Blacks (adjusted odds ratio (AOR): 2.71; 95% confidence interval (CI): 1.36,5.42) and those reporting manic episodes (AOR: 2.41; 95% CI: 1.26,4.60) had higher odds of safe sex counseling. Neither race/ethnicity nor any indicator of mental illness was significantly associated with practicing safer sex because of counseling.

Conclusions: Those with past manic episodes reported more safe sex counseling, which is appropriate given that hypersexuality is a known symptom of mania. Black patients reported more safe sex counseling than white patients, despite controlling for sexual risk. One potential explanation is that counseling was conducted based on assumptions about sexual risk behaviors and patient race. There were no significant disparities in self-reported safer sex practices because of counseling, suggesting that increased counseling did not differentially affect safe sex behavior for black patients and those with manic episodes. Exploring the basis of how patient characteristics can influence counseling and resultant behavior merits further exploration to help reduce disparities in safe sex counseling and outcomes.

Trial registration: NCT00278447

Keywords: Counseling, Disparities, Sexual behavior, Stereotyping

* Correspondence: meredith.damore@bmc.org
[1]Health/care Disparities Research Program, Section of General Internal Medicine, Boston University School of Medicine, Boston Medical Center, 801 Massachusetts Avenue, #2098, Boston, MA 02118, USA
Full list of author information is available at the end of the article

Background

Sexually transmitted infections (STIs) and unintended pregnancy are prevalent among racial/ethnic minorities and individuals with mental illness, and associated with a host of negative health outcomes and costs [1-7]. Unintended pregnancy is costly on many levels, including direct medical costs of births, abortions and fetal losses, indirect costs of wages lost from not working and psychological costs associated with the challenges posed by unintended pregnancy [8-10]. Unintended pregnancies were recently estimated to cost taxpayers $11 billion each year [11,12]. Unintended pregnancies affect the parents (who may endure financial hardship and limitations to their educational attainment) [9,10] and children, in terms of birth outcomes and worse cognitive, emotional and behavioral development [9,10,13-15]. Sexually transmitted infections, including chlamydia, gonorrhea, human papiloma virus (HPV) and the human immunodeficiency virus (HIV), are highly prevalent, preventable infections that can lead to serious health consequences including chronic pain, infertility and mortality [1]. The costs associated with the treatment of STIs are substantial, with an estimated $6.5 billion expended in 2000, for 15-24 year olds alone [16].

In 2001, the rate of unintended pregnancy was highest for black and Hispanic women (98 and 78 per 1,000, respectively), compared with that for whites (35 per 1,000) [3]. More recent data from 2006-2010 report the following percentages of unintended pregnancies resulting in births: 20% non-Hispanic white, 35% Hispanic and 45% black women [17]. Disparities also exist in rates of STIs. For example, black men and women are most affected by chlamydia and gonorrhea, having 9-19 times higher rates than whites [1]. Patients with mental illness have worse health outcomes and a higher medical burden compared to the general population [18-20], including high rates of unintended pregnancy and abortion [5,21,22]. Individuals with mental illness are at higher risk for acquiring STIs and having unintended pregnancies due to increased rates of unprotected intercourse as a result of having less knowledge about contraception, lower capacity to plan ahead, inability to navigate contraceptive resources and being at higher risk for sexual coercion [4-6,23-32].

Substance use, the consumption of alcohol and/or use of illicit drugs, is also associated with sexual risk behaviors [7,33]. The precise link between substance use and engagement in sexual risk behaviors has not been fully established in the literature, but studies suggest that this relationship is primarily a function of less consistent condom use and having multiple sex partners [7,34,35]. Thus, individuals with substance dependence are another vulnerable population at increased risk for unprotected sex, unintended pregnancy and STIs [2].

Identifying barriers to contraceptive usage and safe sexual behaviors is vital in order to prevent STIs and unintended pregnancy. Clinicians can play a pivotal role in educating patients about safe sex and helping to increase their knowledge and use of contraception [36-39]. Although many factors outside of the medical encounter can influence patient sexual behavior and contraception use, clinicians have an opportunity to counsel populations at risk of negative sexual health outcomes, with the potential of affecting patient behavior. This counseling has the potential to reduce risky sexual behaviors and related negative health outcomes [36-41]. Contraceptive counseling in primary care, where there is a preventive focus and a longitudinal relationship with patients, can impact patient contraceptive use and method choice [42].

Despite a higher prevalence of sexual risk behaviors among people with substance dependence than in the general population [7], little is known about disparities in sexual health care among these individuals. Communication problems in the patient-clinician exchange may occur differentially across race/ethnicities and among patients with mental illness, thereby contributing to the disparities in contraceptive use and, subsequently, STIs and unintended pregnancy [37,43]. Primary care clinicians in particular may feel discomfort in discussing these issues because of their sensitive nature and lack of training [44-46]. Despite the vast literature about racial disparities in medical care, little is known about disparities specific to sexual health care [47]. One recent study of low income women found that blacks were more likely to report being pressured by their clinician to use contraceptives, compared to whites [47]. A series of studies found that many African Americans who received family planning care felt discriminated against and held conspiracy beliefs about birth control (such as "birth control is a form of Black genocide") [48].

There is also a dearth of literature assessing disparities in clinician safe sex counseling for people with mental illness [49]. Clinicians may unknowingly make assumptions about the sexual risks of patients with mental illness or their ability to comply with birth control regimens, which may affect clinicians' decisions to provide appropriate counseling and discuss contraceptive options. Sexual health counseling for patients with mental illness is complex and challenging. Clinicians have the challenge of assessing the patient's autonomy in decision-making and risk of unintended pregnancy and STIs. They also need to use judgment to consider how capable the patient is of using contraceptives consistently and effectively [29]. They need to ensure patients fully understand contraceptive options, risks, and benefits [29]. Guidelines to help clinicians assess patient autonomy may be subverted by subconscious bias in decision-making. The role clinicians can play in addressing the sexual health needs of individuals with mental illness warrants further study [50].

Further investigation into the role of clinicians is essential to improving disparities in patients' safe sex behaviors, including the usage of and adherence to contraception. The objectives of this study are therefore to examine whether race/ethnicity and indicators of mental illness are associated with two separate outcomes: patient report of primary clinician's safe sex counseling and practicing safe sex due to counseling among individuals with substance dependence. We hypothesized that minority patients and those with serious mental illness symptoms may receive less safe sex counseling, but anticipate that these disparities will be attenuated after controlling for other patient sociodemographic characteristics, sexual risk behavior and the quality of clinician-patient relationship.

Methods

This was a secondary analysis of cross-sectional data collected for a randomized controlled trial (Addiction Health Evaluation and Disease management (AHEAD) Study) conducted in Boston, Massachusetts from September 2006 to September 2008. This study was approved by the Boston University Medical Campus Institutional Review Board (H-23464). All subjects provided informed consent, and procedures were followed in accordance with the Helsinki Declaration of 1975. A certificate of confidentiality was obtained from the National Institute on Alcohol Abuse and Alcoholism to further protect participants' privacy. The AHEAD study is a randomized clinical trial evaluating chronic disease management for substance dependence in primary care. All subjects had current alcohol and/or drug dependence, by DSM IV criteria [51] (assessed using the Composite International Diagnostic Interview Short Form) [52], were willing to establish or continue primary care at the study location and had engaged in recent heavy drinking or recent drug use. If they had primary care elsewhere but wanted to change to Boston Medical Center (BMC), they were considered eligible. If they had no primary care clinician, they had to be willing to be referred to one at BMC. These subjects were primarily recruited from a residential detoxification unit, but also from a large urban safety-net primary care clinic and through recruitment advertisements on public transportation. Subjects were at least 18 years of age, spoke English or Spanish, and were without indication of cognitive impairment at screening (assessed by a Mini Mental State Examination score greater than 20) [53]. Half of the enrolled subjects were randomized to the AHEAD clinic intervention, which included a team comprised of a nurse care manager, social worker, psychiatrist and internist. All study subjects were reimbursed $35 for completing all baseline visit procedures and $50 at the three month follow-up visit.

The inclusion criteria for our analysis included: completing the 3-month follow-up visit, reported having one particular doctor that they considered to be their regular primary care doctor and being self-reported non-Hispanic black, non-Hispanic white or Hispanic. The question about having a regular doctor asked: "Is there one particular doctor (or primary care provider, e.g. Nurse Practitioner or Physician's Assistant) that you consider to be your regular personal primary care doctor?" Because this survey then continued to use the term "doctor" to encompass all of the primary care clinicians listed above, we continue to use the term "doctor" in reporting the results.

Independent variables

The key independent variables in this study are race/ethnicity and three indicators of serious mental illness. Race/ethnicity was self-reported and categorized as: non-Hispanic white, non-Hispanic black, and Hispanic at the initial study visit. The number of subjects in other racial categories was small (n=20) and therefore were excluded from analyses. The three mental illness variables in this analysis include: moderately severe to severe depression; any past manic episodes; and ≥3 psychotic symptoms, all assessed at the three month study visit. (Mental illness was captured at baseline, but because most participants were starting a detoxification program the 3 month visit provided more accurate data.) Depression was assessed using the Patient Health Questionnaire short form (PHQ-9) which is comprised of nine items about respondents feelings in the last two weeks, such as "feeling down, depressed or hopeless", with responses ranging on a four point scale from "not at all" to "nearly every day" (scores 0-3 points) [54]. We considered someone to be depressed if s/he had a PHQ-9 score of 15 or more, indicating moderately severe to severe depression [54,55]. Past manic or hypomanic episodes were assessed using the Mini International Neuropsychiatric Interview (MINI) [56]. Symptoms of mania include hypersexuality and impulsivity, and thus create the potential for increased sexual risk behavior [57]. This measure used a series of items to assess hypo/mania, asking about the frequency, duration and characteristics of manic episodes, with dichotomous responses of "Yes" or "No". The MINI has been validated as a diagnostic tool in accordance with criteria from to the DSM-IV [56]. Psychotic symptoms were measured using four items from the Behavior and Symptom Identification Scale (BASIS) [58]. These items included: thinking you had special powers, hearing voices or seeing things, thinking people were watching you and thinking people were against you. Subjects were asked about their experience of these symptoms during the past week and asked to rate the frequency of these experiences on a scale ranging from "never" to "always". We considered an individual to have substantial current psychotic symptoms if s/he responded "sometimes", "often", or

"always" on three out of the four items. The complete BASIS measure is comprised of 24 items and using three out of four items not considered diagnostic, but an indication of psychosis.

Dependent variables

The key dependent variables were two items about safe sex counseling and self-attributed resultant patient safe sex behavior, assessed during the three month study interview. The first item, having ever talked about safe sex with your regular doctor, was taken from the Primary Care Assessment Survey (PCAS) [59,60]. The PCAS is a patient reported instrument created to assess several domains that constitute quality primary care [59]. The specific question asked was "Which of the following has your regular doctor ever talked to you about:. . ..safe sex?" (response: yes/no). The second dependent variable, having ever practiced safer sex because of your doctor's advice, was taken from a survey used to examine the relationship between patient income and physician counseling about health risk behaviors [61]. The specific question asked was "Which of the following have you ever done because of your doctor's advice?. . .Practiced safer sex" (response: yes/no). We included all patients who responded to these questions in our analysis. For the second dependent variable, having practiced safer sex because of a doctor's advice, the sample was restricted to only those who had reported ever receiving safe sex counseling.

Covariates

The covariates included in the analyses were self-reported age, gender, education (less than high school; high school; more than high school) and which randomized group the subject was assigned to, taken from the initial study visit. We also included whether the subject had multiple (i.e., >1) male and/or female sex partners in the last three months, as an indication of sexual risk behavior. These data were taken from the audio computer assisted self-interviewing (ACASI) portion of the three month follow-up interview. We included these covariates given their potential effect on receipt of safe sex counseling and behavior [62-64]. We also included quality of the patient-doctor relationship, specifically assessing trust and communication from the PCAS, to evaluate if these variables attenuated any observed relationships between race/ethnicity, mental illness, and the outcomes of interest. The PCAS trust scale was scored based on a series of eight items and the communication scale was based on seven items [59]. These scales were scored as continuous variables and transformed to a scale of 0-100 for multivariable analyses [59].

Statistical analysis

First, descriptive statistics were obtained for all variables stratified by each dependent variable (ever received safe sex counseling and ever practiced safer sex because of this advice). Bivariable tests were also performed for descriptive purposes. Next, we performed a series of multiple logistic regression models to test associations between indicators of mental illness and race/ethnicity with each outcome. For the outcome, having practiced safer sex because of a doctor's advice, the analysis was restricted to the subset who reported ever receiving safe sex counseling. Spearman correlations were used to evaluate potential collinearity between independent variables and covariates. No pair of variables included in the same regression model was highly correlated ($r>0.40$). The following four models were fit for each of the two dependent variables. Model 1 was a preliminary, minimally adjusted model that included the main independent variables race/ethnicity and indicators of mental illness (depression, having had any manic episodes and psychotic symptoms), and two potential confounders: randomization group and multiple sex partners in the last three months. Model 2 additionally controlled for key sociodemographic characteristics: age, education, and gender. Model 3 added the trust scale from the PCAS as a continuous variable. Model 4, the final model representing the primary analyses, included the communication scale from PCAS as a continuous variable, and removed the trust scale. We did not include trust and communication scores in the same model due to their high correlation ($r=0.67$). The findings reported in the Results section are taken from the final model (Model 4) with communication score, unless otherwise specified. Adjusted odds ratios (AOR) and 95% confidence intervals (CI) are reported. All analyses were conducted using two-sided tests and a significance level of 0.05. Due to the exploratory nature of the analyses, adjustments were not made for multiple comparisons. Statistical analyses were performed by the Boston University School of Public Health's Data Coordinating Center using SAS software (version 9.1; SAS Institute, Cary, NC).

Results

Of the 563 individuals enrolled in the AHEAD study, 500 were followed up at 3 months. Of those, 299 reported that they had a regular doctor and 295 of those answered the two questions about safe sex counseling and behavior. After excluding the 20 subjects who were not black, white or Hispanic, our final sample for the analysis of ever received safe sex counseling included 275 individuals. The study sample was comprised of 44% white, 45% black and 11% Hispanic women and men (Table 1), with a mean age of 40 (range 18-61) and median age of 42 (not shown). Moderately severe to severe depression was common, reported among 58% of subjects. Forty percent reported a previous manic episode and 16% had at least three psychotic symptoms. Approximately 71% (195/275) of the sample reported having a doctor ever talk to them about safe sex. These 195

Table 1 Sample characteristics by receipt of safe sex counseling and practicing safer sex due to counseling

	Doctor ever talked about safe sex n (%)				Ever practiced safe sex because of doctor's advice n (%)			
	Yes (n=195)	No (n=80)	Total[a] (n=275)	p-value [b]	Yes (n=149)	No (n=46)	Total[a] (n=195)	p-value [b]
Race/ethnicity								
Non-Hispanic white	74 (61.2)	47 (38.8)	121	**0.007**	51 (68.9)	23 (31.1)	74	0.12
Non-Hispanic black	97 (78.9)	26 (21.1)	123		77 (79.4)	20 (20.6)	97	
Hispanic	24 (77.4)	7 (22.6)	31		21 (87.5)	3 (12.5)	24	
Moderately severe to Severe depression (PHQ-9 score of ≥ 15)								
Yes	114 (71.7)	45 (28.3)	159	0.82	89 (78.1)	25 (21.9)	114	0.52
No	81 (70.4)	34 (29.6)	115		60 (74.1)	21 (25.9)	81	
Any manic or hypomanic episode								
Yes	88 (80.0)	22 (20.0)	110	**0.007**	67 (76.1)	21 (23.9)	88	0.94
No	107 (64.8)	58 (35.2)	165		82 (76.6)	25 (23.4)	107	
Three or more psychotic symptoms[c]								
Yes	36 (83.7)	7 (16.3)	43	**0.04**	31 (86.1)	5 (13.9)	36	0.19
No	159 (68.5)	73 (31.5)	232		118 (74.2)	41 (25.8)	159	
Age (years)								
<30	37 (69.8)	16 (30.2)	53	0.82	27 (73.0)	10 (27.0)	37	0.3
31-39	43 (72.9)	16 (27.1)	59		32 (74.4)	11 (25.6)	43	
40-49	80 (72.7)	30 (27.3)	110		59 (73.8)	21 (26.2)	80	
≥50	35 (66.0)	18 (34.0)	53		31 (88.6)	4 (11.4)	35	
Gender								
Female	70 (79.6)	18 (20.4)	88	**0.03**	55 (78.6)	15 (21.4)	70	0.6
Male	125 (66.8)	62 (33.2)	187		94 (75.2)	31 (24.8)	125	
Education								
Less than high school	49 (79.0)	13 (21.0)	62	0.28	39 (79.6)	10 (20.4)	49	0.82
High school graduate	91 (68.4)	42 (31.6)	133		69 (75.8)	22 (24.2)	91	
More than high school	55 (68.8)	25 (31.2)	80		41 (74.6)	14 (25.4)	55	
Randomization group								
Control	102 (72.3)	39 (27.7)	141	0.59	78 (76.5)	24 (23.5)	102	0.98
Intervention	93 (69.4)	41 (30.6)	134		71 (76.3)	22 (23.7)	93	
Multiple Sex Partners [d]								
No	134 (71.7)	53 (28.3)	187	0.58	108 (80.6)	26 (19.4)	134	**0.03**
Yes	56 (68.3)	26 (31.7)	82		37 (66.1)	19 (33.9)	56	
PCAS trust score								
0-25	2 (66.7)	1 (33.3)	3	**0.02**	2 (100.0)	0 (0.0)	2	0.83
26-50	11 (45.8)	13 (54.2)	24		8 (72.7)	3 (27.3)	11	
51-75	91 (70.0)	39 (30.0)	130		67 (73.6)	24 (26.4)	91	
76-100	91 (77.1)	27 (22.9)	118		72 (79.1)	19 (20.9)	91	
PCAS communication score								
0-25	1 (33.3)	2 (66.7)	3	0.06	1 (100.0)	0 (0.0)	1	0.9
26-50	16 (57.1)	12 (42.9)	28		12 (75.0)	4 (25.0)	16	
51-75	68 (68.7)	31 (31.3)	99		51 (75.0)	17 (25.0)	68	
76-100	110 (76.4)	34 (23.6)	144		85 (77.3)	25 (22.7)	110	

[a] missing responses not shown.
[b] p-values calculated based on chi-square or Fisher's exact tests as appropriate; p-values in bold are statistically significant (p<0.05).
[c] Defined as responding "sometimes", "often" or "always" on three out of four BASIS24 questions.
[d] >1 sex partner within the past three months.

subjects comprised the sample for the analysis of ever practiced safer sex because of your doctor's advice. Among the subset who received counseling, 76% (149/195) reported practicing safe sex because of their doctor's advice. In the adjusted models, higher trust (AOR=1.03, 95% CI: 1.01, 1.04; Model 3) and better communication (AOR=1.02, 95% CI: 1.01, 1.04) were associated with higher odds of receipt of safe sex counseling (per 1 point increase on a scale of 0-100; Table 2). Having multiple sex partners in the last three months (AOR=0.44, 95% CI: 0.21, 0.94) was associated with lower odds of practicing safer sex because of this advice, compared to those with 0-1 sex partners (Table 3).

Results by race/ethnicity

In bivariable analyses, whites had the lowest observed proportion of their doctor ever talking to them about safe sex (p=0.007; Table 1), compared to blacks or Hispanics. There were no significant findings by race/ethnicity for practicing safer sex because of a doctor's advice. In the multivariable models, after adjustment for covariates, black subjects had significantly higher odds of a doctor having ever talked with them about safe sex (AOR=2.71, 95% CI: 1.36, 5.42; Table 2) compared to whites. No racial/ethnic group was associated with significantly higher odds of reporting ever practicing safe sex because of a doctor's advice compared to whites (Table 3). Results were similar across all adjusted models.

Results by indicators of mental illness

In bivariable analyses, those who reported a manic episode (p=0.007) or three or more psychotic symptoms (p=0.04) had higher odds of reporting that their doctor had ever talked to them about safe sex (Table 1), compared to those who had not had an episode or had less than three psychotic symptoms, respectively. None of

Table 2 Multivariable logistic regression models for receipt of safe sex counseling from primary care doctor

	Doctor ever talked about safe sex			
	Model 1	**Model 2**	**Model 3**	**Model 4**
	AOR (95% Confidence interval)			
Main Independent Variables				
Race/ethnicity				
Non-Hispanic black vs. white	**2.38 (1.28, 4.44)** *	**2.82 (1.43, 5.56)** *	**2.78 (1.40, 5.56)** **	**2.71 (1.36, 5.42)** *
Hispanic vs. white	1.87 (0.72, 4.87)	2.12 (0.77, 5.81)	2.58 (0.91, 7.26)	2.36 (0.83, 6.66)
Moderately severe to Severe depression (PHQ-9 score of ≥ 15)				
Yes vs. No	1.05 (0.59, 1.89)	1.02 (0.56, 1.86)	1.05 (0.57, 1.94)	0.98 (0.53, 1.82)
Any manic or hypomanic episode				
Yes vs. No	**2.06 (1.12, 3.78)** *	**2.00 (1.08, 3.72)** *	**2.38 (1.25, 4.56)** **	**2.41 (1.26, 4.60)** **
Three or more psychotic symptoms [a]				
Yes vs. No	1.56 (0.62, 3.88)	1.80 (0.70, 4.58)	1.78 (0.68, 4.62)	1.78 (0.69, 4.62)
Covariates				
Randomization group				
Intervention vs. Control	0.86 (0.49, 1.49)	0.84 (0.48, 1.50)	0.87 (0.48, 1.57)	0.78 (0.43, 1.41)
Multiple Sex Partners (past 3 months)				
Yes vs. No	0.67 (0.37, 1.23)	0.65 (0.35, 1.22)	0.77 (0.41, 1.46)	0.77 (0.41, 1.45)
Age [b]		0.97 (0.94, 1.00)	0.97 (0.94, 1.00)	0.97 (0.94, 1.00)
Gender				
Female vs. Male		1.85 (0.97, 3.55)	1.70 (0.87, 3.30)	1.74 (0.90, 3.38)
Education				
Less than high school vs. More than high school		1.24 (0.54, 2.85)	1.32 (0.56, 3.10)	1.45 (0.61, 3.46)
High school graduate vs. More than high school		0.85 (0.44, 1.63)	0.91 (0.47, 1.77)	0.92 (0.48, 1.79)
PCAS trust score [c]			**1.03 (1.01, 1.04)** **	
PCAS communication score [c]				**1.02 (1.01, 1.04)** **

* p<0.05.
** p<0.01.
[a] Defined as responding "sometimes", "often" or "always" on three out of four BASIS24 questions.
[b] AORs associated with a 1 year increase in age.
[c] AORs associated with a 1 point increase in PCAS scale score.

Table 3 Multivariable logistic regression models for practicing safer sex due to counseling from primary care doctor

	Ever practiced safe sex because of doctor's advice			
	Model 1	Model 2	Model 3	Model 4
	AOR (95% Confidence interval)			
Main Independent Variables				
Race/ethnicity				
Non-Hispanic black vs. white	1.87 (0.88, 3.97)	1.84 (0.80, 4.22)	1.85 (0.81, 4.25)	1.85 (0.81, 4.26)
Hispanic vs. white	2.48 (0.65, 9.53)	2.56 (0.64, 10.19)	2.49 (0.63, 9.95)	2.56 (0.64, 10.17)
Moderately severe to Severe depression (PHQ-9 score of ≥ 15)				
Yes vs. No	1.26 (0.60, 2.65)	1.26 (0.60, 2.67)	1.25 (0.59, 2.65)	1.27 (0.60, 2.69)
Any manic or hypomanic episode				
Yes vs. No	0.84 (0.41, 1.74)	0.86 (0.41, 1.78)	0.82 (0.38, 1.74)	0.84 (0.39, 1.78)
Three or more psychotic symptoms [a]				
Yes vs. No	2.05 (0.70, 6.07)	2.05 (0.68, 6.16)	2.06 (0.68, 6.18)	2.05 (0.68, 6.15)
Covariates				
Randomization group				
Intervention vs. Control	1.01 (0.50, 2.03)	1.01 (0.50, 2.06)	1.00 (0.49, 2.04)	1.01 (0.50, 2.06)
Multiple Sex Partners (past 3 months)				
Yes vs. No	**0.43 (0.20, 0.89)** *	**0.44 (0.21, 0.94)** *	**0.44 (0.20, 0.93)** *	**0.44 (0.21, 0.94)** *
Age [b]		1.00 (0.96, 1.04)	1.00 (0.96, 1.04)	1.00 (0.96, 1.04)
Gender				
Female vs. Male		1.26 (0.58, 2.73)	1.31 (0.60, 2.88)	1.29 (0.59, 2.82)
Education				
Less than high school vs. More than high school		1.18 (0.43, 3.25)	1.15 (0.41, 3.18)	1.17 (0.42, 3.23)
High school graduate vs. More than high school		0.95 (0.41, 2.23)	0.92 (0.39, 2.18)	0.93 (0.39, 2.22)
PCAS trust score [c]			0.99 (0.97, 1.02)	
PCAS communication score [c]				1.00 (0.98, 1.02)

* $p<0.05$.
[a] Defined as responding "sometimes", "often" or "always" on three out of four BASIS24 questions.
[b] AORs associated with a 1 year increase in age.
[c] AORs associated with a 1 point increase in PCAS scale score.

the mental illness variables were significantly related to practicing safer sex because of a doctor's advice. In fully adjusted models, having a manic episode remained significantly associated with higher odds of a doctor having ever talked to them about safe sex (AOR=2.41, 95% CI: 1.26, 4.60; Table 2). The association was statistically significant across all adjusted models. None of the mental health variables were significantly associated with practicing safer sex because of their doctor's advice in adjusted models (Table 3).

Discussion

The majority of adults with substance dependence in our sample reported being counseled about safe sex from their primary care doctor. Of those who were counseled, about three quarters reported that they practiced safer sex because of this advice. Our results suggest that improvements in safe sex counseling are needed for all individuals with substance dependence, particularly whites. This is consistent with prior literature that suggests general deficiencies in safe sex counseling and the need to integrate this into regular health care visits [41,61,65]. Lack of counseling may be a result of primary care clinicians feeling it is not their responsibility to counsel patients in sexual health behaviors [44]. Clinicians may also be uncomfortable in sexual health counseling and need more training related to sexual health [46]. Time constraints and lack of support staff have been cited as potential reasons for lack of sexual health counseling [44,45,66]. The expanded role of nurses or case managers may be a potential solution in improving quality sexual health care [45]. Of note, patient trust and communication with their doctor were the only covariates associated with receipt of safe sex counseling. This suggests that improving patient-doctor trust and communication may increase safe sex counseling for patients with substance dependence.

Our findings suggest that safe sex counseling can have a positive impact on patient behaviors. Understanding

the discourse between patients and clinicians and effective methods for delivering contraceptive information is critical in developing practice guidelines and guidance for clinicians [67]. Future research should focus on understanding how clinicians can affect patient behavior, in an effort to reduce disparities in sexual health and outcomes.

Blacks and those with manic episodes reported being counseled by their doctors significantly more about safe sex. Although most who received counseling reported practicing safe sex because of their doctor's advice, this was not more likely to translate into safer sex practices for one racial/ethnic group or individuals with indicators of mental illness. Interestingly, the covariates included in both sets of analyses largely did not affect the significant associations between race/ethnicity or mental illness with our outcomes. Although the findings that blacks are counseled more about safe sex are in contrast to our hypotheses, there is literature to support the finding of increased counseling for minorities [62]. Previous disparities research suggests that clinician behavior may be influenced by population based studies that affect their decision making in counseling patients [68,69]. Thus, if clinicians know from the literature that racial/ethnic minorities are more likely to have negative sexual health outcomes, they may counsel these patients more about safe sex practices. Clinicians may also have implicit, sub-conscious biases related to race that are not fully apparent but could affect patient counseling [70,71]. Counseling based on assumptions about sexual risk behaviors because of patient race can be considered stereotyping, that leads to disparities in counseling black patients more about safe sex. Instead, patient-centered counseling, based on individual risk factors, tailored to that individual's needs, is ideal.

Patients with manic episodes may have been counseled more, given that hypersexuality is associated with these episodes [57]. It is difficult to disentangle the meaning of this finding, given that the subject was asked about current psychotic symptoms and depression, but *any* prior history of manic episodes. It is not clear if the manifestation of symptoms or diagnosis of mental illness came before or after safe sex counseling and thus we cannot determine if increased counseling among this group is in fact related to their mental illness. We also cannot verify that the clinician knew of the patient's mental illness.

Although many interventions for increasing safe sex practices can be found in the literature [72], particularly for young and poor, minority women, there is not a major focus on individuals with substance dependence. Most of the interventions targeting substance users focus solely on reducing HIV infection [62,73], instead of prevention of general STIs and unintended pregnancy. Interventions should focus on the interaction of factors, such as substance use, race/ethnicity and mental illness, that may lead to disparities in sexual health outcomes [63]. Ultimately, reducing negative sexual health outcomes will require a multipronged intervention targeted at the patient, clinician and system level to help influence patient behavior and enhance the patient-clinician interaction [37].

Our findings should be interpreted within the limitations of our study. Although our analysis examined three racial/ethnic groups, we did not test for within group differences (i.e., by type of Hispanic origin) or interactions by gender, because of potential sample size issues [74]. The findings that black patients and those with manic episodes had significantly higher odds of reporting being counseled could also be due to differential response bias or incomplete statistical adjustment (i.e., history of sexually transmitted infections). This study is based entirely on patient self-report, which may be subject to recall or social desirability bias. However, understanding if safe sex counseling occurred and how it affected behavior from patient's perspective is ideal because they can most accurately report if/how counseling influenced their practices. Our sample was enrolled within one metropolitan area and the majority of subjects was of low socioeconomic status and thus may not be generalizable to all other substance dependent populations. We were unable to determine if patients received safe sex counseling from another type of clinician (other than their primary care doctor) or practiced safer sex because of advice from another clinician. There may also be temporal issues, in asking about lifetime safe sex counseling versus symptoms of mental illnesses, which may have manifested after counseling occurred (as described above).

We did not have data to account for the U.S. Preventive Services Task Force recommendations of counseling if patients had an STI within the last year, which may have affected safe sex counseling. Instead, we used number of sex partners in the last three months as an indicator of sexual risk behavior. We also used only partial diagnostic instruments as indicators of mental illness, particularly for psychotic symptoms. Despite these limitations, this study is among the first to examine racial/ethnic and mental health disparities in safe sex counseling and practices among individuals with substance dependence. This paper is one of many that will answer the call for studies to test hypotheses about how clinician behaviors affect disparities in health care [68]. Beyond merely documenting that counseling occurred, research needs to focus on enhancing our understanding of how counseling can have an impact on patient risk behavior.

Conclusions

Black patients and those with a history of manic episodes were found to be counseled more about safe sex than white patients and those without manic episodes, respectively, despite controlling for sexual risk. Exploring the basis of

how patient characteristics can influence counseling and resultant behavior merits further exploration to help reduce disparities in safe sex counseling and outcomes.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors have made substantial contributions to the study's conception, design, data collection, analysis and/or interpretation of results, been involved in drafting or revising the manuscript and have approved the version to be published. *MMD* conceived of this study's analysis plan, interpreted the results and drafted and revised the manuscript. *DMC* is a biostatistician and provided guidance in developing and revising the analytic plan, interpreting the results and revising the manuscript. *DAD* conducted the statistical programming for this study and revised the manuscript for publication. *JHS* was Principal Investigator of one of the grants that provided support for this study, made substantial contributions to the conception, design, data collection, analytic plan and data interpretation for this study and also repeatedly revised the manuscript. *RS* was Principal Investigator of one of the grants that provided support for this study, made substantial contributions to the conception, design, data collection, analysis and interpretation for this study and made multiple revisions to this manuscript.

Acknowledgements
This project was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) Grant Number R01AA010870, National Institute on Drug Abuse (NIDA) Grant Number R01DA010019, and National Center for Research Resources (NCRR) Grant Number UL1RR025771. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIAAA, NIDA, NCRR or the National Institutes of Health. The authors would like to acknowledge Victoria Parker, DBA, Lois McCloskey, DrPH, and Barbara Bokhour, PhD (Boston University School of Public Health) for their help in editing this manuscript.

Author details
[1]Health/care Disparities Research Program, Section of General Internal Medicine, Boston University School of Medicine, Boston Medical Center, 801 Massachusetts Avenue, #2098, Boston, MA 02118, USA. [2]Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA 02118, USA. [3]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA. [4]School of Health Sciences, Cleveland State University, Cleveland, OH 44115, USA. [5]Department of Community Health Sciences, Boston University School of Public Health, Boston, MA 02118, USA. [6]Department of Epidemiology, Boston University School of Public Health, Boston, MA 02118, USA.

Received: 20 September 2012 Accepted: 21 December 2012
Published: 31 December 2012

References
1. *Facts on sexually transmitted infections in the united states*. New York, NY: Guttmacher Institute; 2009.
2. Substance Abuse and Mental Health Services Administration, Office of Applied Studies: *The NSDUH Report: Sexually Transmitted Diseases and Substance Use*. Rockville, MD; 2007.
3. Finer L, Henshaw S: **Disparities in rates of unintended pregnancy in the United States, 1994-2001.** *Perspect Sex Reprod Health* 2006, **38:**90–96.
4. Meade CS, Sikkema KJ: **HIV risk behavior among adults with severe mental illness: a systematic review.** *Clin Psychol Rev* 2005, **25**(4):433–457.
5. Miller LJ: **Sexuality, reproduction, and family planning in women with schizophrenia.** *Schizophr Bull* 1997, **23**(4):623–635.
6. Rosenberg SD, *et al*: **Prevalence of HIV, hepatitis B, and hepatitis C in people with severe mental illness.** *Am J Public Health* 2001, **91**(1):31–37.
7. *Substance use and risky sexual activity*. Washington, DC: Kaiser Family Foundation; 2002.
8. Trussell J: **The cost of unintended pregnancy in the United States.** *Contraception* 2007, **75:**168–170.
9. Brown S, Eisenberg L: *The best intentions: unintended pregnancy and the well-being of children and families*. Washington, DC: National Academy Press; 1995:80–82.
10. Logan C, *et al*: *The consequences of unintended childbearing: a white paper*. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy Child Trends, Inc; 2007.
11. Sonfield A, *et al*: **The public costs of births resulting from unintended pregnancies: national and state-level estimates.** *Perspect Sex Reprod Health* 2011, **43**(2):94–102.
12. Monea E, Thomas A: **Unintended pregnancy and taxpayer spending.** *Perspect Sex Reprod Health* 2011, **43**(2):88–93.
13. Santelli J, *et al*: **The measurement and meaning of unintended pregnancy.** *Perspect Sex Reprod Health* 2003, **35**(2):94–101.
14. Sable M, *et al*: **Pregnancy wantedness and adverse pregnancy outcomes: differences by race and Medicaid status.** *Fam Plann Perspect* 1997, **29**(2):76–81.
15. Mohllajee A, *et al*: **Pregnancy intention and its relationship to birth and maternal outcomes.** *Obstet Gynecol* 2007, **109**(3):678–686.
16. Chesson HW, *et al*: **The estimated direct medical cost of sexually transmitted diseases among American youth, 2000.** *Perspect Sex Reprod Health* 2004, **36**(1):11–19.
17. Mosher WD, Jones J, Abma JC: **Intended and unintended births in the United States: 1982-2010.** *Natl Health Stat Rep* 2012, **55:**1–28.
18. Carney CP, Jones LE: **Medical comorbidity in women and men with bipolar disorders: a population-based controlled study.** *Psychosom Med* 2006, **68**(5):684–691.
19. Carney CP, Jones L, Woolson RF: **Medical comorbidity in women and men with schizophrenia: a population-based controlled study.** *J Gen Intern Med* 2006, **21**(11):1133–1137.
20. Soreca I, *et al*: **Relationship of general medical burden, duration of illness and age in patients with bipolar I disorder.** *J Psychiatr Res* 2008, **42**(11):956–961.
21. Miller LJ, Finnerty M: **Sexuality, pregnancy, and childrearing among women with schizophrenia-spectrum disorders.** *Psychiatr Serv* 1996, **47**(5):502–506.
22. Barkla J, *et al*: **Pregnancy in women with psychotic disorders.** *Arch Womens Ment Health* 2000, **3:**23–26.
23. Miller LJ, Finnerty M: **Family planning knowledge, attitudes and practices in women with schizophrenic spectrum disorders.** *J Psychosom Obstet Gynaecol* 1998, **19**(4):210–217.
24. Raja M, Azzoni A: **Sexual behavior and sexual problems among patients with severe chronic psychoses.** *Eur Psychiatry* 2003, **18**(2):70–76.
25. Egan TM, Siegert RJ, Fairley NA: **Use of hormonal contraceptives in an institutional setting: reasons for use, consent and safety in women with psychiatric and intellectual disabilities.** *N Z Med J* 1993, **106**(961):338–341.
26. Leutar Z, Mihokovic M: **Level of knowledge about sexuality of people with mental disabilities.** *Sex Disabil* 2007, **25**(3):93–109.
27. Cook J: **Sexuality and people with psychiatric disabilities.** *Sex Disabil* 2000, **18**(3):195–206.
28. Weinhardt L, Bickham N, Carey M: **Sexual coercion among women living with a severe and persistent mental illness: review of the literature and recommendations for mental health providers.** *Aggression Violent Behav* 1999, **4**(3):307–317.
29. Neinstein LS, Katz B: **Psychiatrically impaired patients often exposed to pregnancy risk.** *Contracept Technol Update* 1985, **6**(6):88–90.
30. Dyer JG, McGuinness TM: **Reducing HIV risk among people with serious mental illness.** *J Psychosoc Nurs Ment Health Serv* 2008, **46**(4):26–34.
31. Beyer JL, *et al*: **Prevalence of HIV infection in a general psychiatric outpatient population.** *Psychosomatics* 2007, **48**(1):31–37.
32. Carey MP, Carey KB, Kalichman SC: **Risk for human immunodeficiency virus (HIV) infection among persons with severe mental illnesses.** *Clin Psychol Rev* 1997, **17**(3):271–291.
33. Saraceno B, Yong Kim J: *Alcohol use and sexual risk behaviour: a cross-cultural study in eight countries*. Geneva, Switzerland: World Health Organization, Department of Mental Health and Substance Abuse; 2005.
34. Pallonen UE, *et al*: **Personal and partner measures in stages of consistent condom use among African-American heterosexual crack cocaine smokers.** *AIDS Care* 2008, **20**(2):205–213.
35. Schilling RF, *et al*: **Sexual behavior, attitudes toward safer sex, and gender among a cohort of 244 recovering i.v. drug users.** *Int J Addict* 1991, **26**(8):859–877.

36. Petersen R, et al: **Pregnancy and STD prevention counseling using an adaptation of motivational interviewing: a randomized controlled trial.** *Perspect Sex Reprod Health* 2007, **39**(1):21–28.
37. Dehlendorf C, et al: **Disparities in family planning.** *Am J Obstet Gynecol* 2010, **202**(3):214–220.
38. Moos MK, et al: **Healthier women, healthier reproductive outcomes: recommendations for the routine care of all women of reproductive age.** *Am J Obstet Gynecol* 2008, **199**(6 Suppl 2):S280–S289.
39. Nobili MP, et al: **The effect of patient-centered contraceptive counseling in women who undergo a voluntary termination of pregnancy.** *Patient Educ Couns* 2007, **65**(3):361–368.
40. Workowski KA, Berman SM: **Sexually transmitted diseases treatment guidelines, 2006.** *MMWR Recomm Rep* 2006, **55**(RR-11):1–94.
41. Tao G, Irwin KL, Kassler WJ: **Missed opportunities to assess sexually transmitted diseases in U.S. adults during routine medical checkups.** *Am J Prev Med* 2000, **18**(2):109–114.
42. Lee JK, et al: **The impact of contraceptive counseling in primary care on contraceptive use.** *J Gen Intern Med* 2011, **26**(7):731–736.
43. Thorburn S, Bogart LM: **African American women and family planning services: perceptions of discrimination.** *Women Health* 2005, **42**(1):23–39.
44. Ashton MR, et al: **Primary care physician attitudes regarding sexually transmitted diseases.** *Sex Transm Dis* 2002, **29**(4):246–251.
45. Gott M, et al: **"Opening a can of worms": GP and practice nurse barriers to talking about sexual health in primary care.** *Fam Pract* 2004, **21**(5):528–536.
46. Khan A, et al: **Does physician bias affect the quality of care they deliver? Evidence in the care of sexually transmitted infections.** *Sex Transm Infect* 2008, **84**(2):150–151.
47. Becker D, Tsui AO: **Reproductive health service preferences and perceptions of quality among low-income women: racial, ethnic and language group differences.** *Perspect Sex Reprod Health* 2008, **40**(4):202–211.
48. Thorburn S, Bogart LM: **Conspiracy beliefs about birth control: barriers to pregnancy prevention among African Americans of reproductive age.** *Health Educ Behav* 2005, **32**(4):474–487.
49. Falvey J: **Clinical judgment in case conceptualization and treatment planning across mental health disparities.** *J Couns Dev* 2001, **79**:292–303.
50. Brown AP, Lubman DI, Paxton SJ: **STIs and blood borne viruses - risk factors for individuals with mental illness.** *Aust Fam Physician* 2008, **37**(7):531–534.
51. American psychiatric association: *Diagnostic and statistical manual of mental disorders. 4th ed.* Washington, DC: American Psychiatric Association; 2000.
52. Gigantesco A, Morosini P: **Development, reliability and factor analysis of a self-administered questionnaire which originates from the world health Organization's composite international diagnostic interview - short form (CIDI-SF) for assessing mental disorders.** *Clin Pract Epidemiol Ment Health* 2008, **4**:8.
53. Klein LE, et al: **Diagnosing dementia. Univariate and multivariate analyses of the mental status examination.** *J Am Geriatr Soc* 1985, **33**(7):483–488.
54. Kroenke K, Spitzer RL, Williams JB: **The PHQ-9: validity of a brief depression severity measure.** *J Gen Intern Med* 2001, **16**(9):606–613.
55. Gilbody S, et al: **Screening for depression in medical settings with the patient health questionnaire (PHQ): a diagnostic meta-analysis.** *J Gen Intern Med* 2007, **22**(11):1596–1602.
56. Sheehan DV, et al: **The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10.** *J Clin Psychiatry* 1998, **59**(Suppl 20):22–33. quiz 34-57.
57. Meade CS, et al: **HIV risk behavior among patients with co-occurring bipolar and substance use disorders: associations with mania and drug abuse.** *Drug Alcohol Depend* 2008, **92**(1–3):296–300.
58. Eisen SV, et al: **Reliability and validity of the BASIS-24 mental health survey for whites, African-americans, and latinos.** *J Behav Health Serv Res* 2006, **33**(3):304–323.
59. Safran DG, et al: **The primary care assessment survey: tests of data quality and measurement performance.** *Med Care* 1998, **36**(5):728–739.
60. Safran DG, et al: **Linking primary care performance to outcomes of care.** *J Fam Pract* 1998, **47**(3):213–220.
61. Taira DA, et al: **The relationship between patient income and physician discussion of health risk behaviors.** *JAMA* 1997, **278**(17):1412–1417.
62. Cao D, Marsh JC, Shin HC: **Gender and racial/ethnic disparities in the impact of HIV prevention programming in substance abuse treatment.** *Am J Drug Alcohol Abuse* 2008, **34**(6):730–740.
63. Kraut-Becher J, et al: **Examining racial disparities in HIV: lessons from sexually transmitted infections research.** *J Acquir Immune Defic Syndr* 2008, **47**(Suppl 1):S20–S27.
64. Lin JS, Whitlock E, O'Connor E, Bauer V: **Behavioral Counseling to Prevent Sexually Transmitted Infections: A Systematic Review for the U.S. Preventive Services Task Force.** *Ann Intern Med* 2008, **149**(7):497–508.
65. Weisman CS, et al: **Contraceptive counseling in managed care: preventing unintended pregnancy in adults.** *Womens Health Issues* 2002, **12**(2):79–95.
66. Mark H, et al: **Providers' perceived barriers to sexually transmitted disease care in 2 large health maintenance organizations.** *Sex Transm Dis* 2008, **35**(2):184–189.
67. Hayter M: **Nurses' discourse in contraceptive prescribing: an analysis using Foucault's 'procedures of exclusion'.** *J Adv Nurs* 2007, **58**(4):358–367.
68. van Ryn M, Fu SS: **Paved with good intentions: do public health and human service providers contribute to racial/ethnic disparities in health?** *Am J Public Health* 2003, **93**(2):248–255.
69. van Ryn M, Burke J: **The effect of patient race and socio-economic status on physicians' perceptions of patients.** *Soc Sci Med* 2000, **50**(6):813–828.
70. Smedley BD, Stith AY, Nelson AR, Unequal Treatment: *Confronting racial and ethnic disparities in health care. Institute of medicine.* Washington, DC: National Academies Press; 2002.
71. Sabin J, et al: **Physicians' implicit and explicit attitudes about race by MD race, ethnicity, and gender.** *J Health Care Poor Underserved* 2009, **20**(3):896–913.
72. Kirby D: **The impact of programs to increase contraceptive use among adult women: a review of experimental and quasi-experimental studies.** *Perspect Sex Reprod Health* 2008, **40**(1):34–41.
73. Mausbach BT, et al: **Efficacy of a behavioral intervention for increasing safer sex behaviors in HIV-negative, heterosexual methamphetamine users: results from the Fast-Lane Study.** *Ann Behav Med* 2007, **34**(3):263–274.
74. Burlew AK, et al: **Measurement and data analysis in research addressing health disparities in substance abuse.** *J Subst Abuse Treat* 2009, **36**(1):25–43.

doi:10.1186/1742-4755-9-35
Cite this article as: D'Amore et al.: Disparities in safe sex counseling & behavior among individuals with substance dependence: a cross-sectional study. *Reproductive Health* 2012 **9**:35.

Effectiveness of a Risk Screener in Identifying Hepatitis C Virus in a Primary Care Setting

Mari-Lynn Drainoni, PhD, Alain H. Litwin, MD, Bryce D. Smith, PhD, Elisa A. Koppelman, MSW, M. Diane McKee, MD, Cindy L. Christiansen, PhD, Allen L. Gifford, MD, Cindy M. Weinbaum, MD, and William N. Southern, MD

Objectives. We evaluated an intervention designed to identify patients at risk for hepatitis C virus (HCV) through a risk screener used by primary care providers.

Methods. A clinical reminder sticker prompted physicians at 3 urban clinics to screen patients for 12 risk factors and order HCV testing if any risks were present. Risk factor data were collected from the sticker; demographic and testing data were extracted from electronic medical records. We used the *t* test, χ^2 test, and rank-sum test to compare patients who had and had not been screened and developed an analytic model to identify the incremental value of each element of the screener.

Results. Among screened patients, 27.8% (n = 902) were identified as having at least 1 risk factor. Of screened patients with risk factors, 55.4% (n = 500) were tested for HCV. Our analysis showed that 7 elements (injection drug use, intranasal drug use, elevated alanine aminotransferase, transfusions before 1992, ≥ 20 lifetime sex partners, maternal HCV, existing liver disease) accounted for all HCV infections identified.

Conclusions. A brief risk screener with a paper-based clinical reminder was effective in increasing HCV testing in a primary care setting. (*Am J Public Health.* 2012;102:e115–e121. doi:10.2105/AJPH.2012.300659)

Hepatitis C virus (HCV) is a significant public health problem. With 3.2 million Americans chronically infected,[1] HCV is the leading cause of liver-related deaths,[2] accounting for 15 000 deaths in 2007.[3] Although earlier treatments were moderately effective in reducing the HCV disease burden,[4] new treatments with greater promise have become available.[5] Because treatment cannot be offered without diagnosis and 45% to 85% of patients with HCV are unaware of their infection,[6,7] interventions designed to increase the number of HCV cases diagnosed are urgently needed.

Guidelines for HCV screening vary. The Centers for Disease Control and Prevention (CDC) recommends that patients who have injected drugs, who have long-term hemodialysis histories or persistently abnormal alanine aminotransferase (ALT) levels, who had blood transfusions or organ transplants before July 1992 (when HCV was eradicated from the nation's blood supply), who have been exposed to HCV (e.g., their mothers were HCV positive or they have been exposed at their workplace), or who are HIV positive[8] be assessed for HCV risk. Other authorities have expanded recommendations to include current sexual partners of individuals with HCV,[9,10] people who have had multiple sex partners, intranasal cocaine users, people with tattoos or repeated body piercings, people with high levels of daily alcohol use over time, Vietnam-era veterans,[11] and immigrants from countries with high HCV prevalence rates.[12]

In addition, with respect to research on HCV risk, various studies have shown that homelessness, incarceration,[13] tattoos,[14] barbershop shaving,[15] body piercing,[16] ear piercing among men,[17] use of intranasal drugs and crack cocaine,[18] and mental illness[19] are associated with higher risk. Although not explicitly recommending testing, this literature suggests that these are potential HCV risk factors for which screening may be appropriate.

Multiple approaches can be used in HCV testing programs. Universal screening of people with identified risks appears to best meet CDC's recommendations and to be the most efficient strategy, given that individuals with identified risk factors have been shown to have a much higher prevalence of HCV than the general population.[1] As the front-line health care providers for most Americans, primary care settings offer an important opportunity to incorporate HCV risk assessments, although examination of this model has been limited.

In 2 studies conducted in primary care settings, patient self-administered questionnaires have been used to assess HCV risk screening. In one of these studies, set in an urban clinic, patients completed a 27-item risk assessment[20]; the other study, set in a Veterans Health Administration facility, involved a retrospective analysis of HCV testing among veterans who had reported HCV risk factors on a self-administered questionnaire.[21] To date, no HCV screening tools have been validated, and no studies comparing different types of interventions have been conducted, including comparisons of patient-completed screening instruments and screeners implemented by primary care providers (PCPs).

We implemented a PCP-based risk screening intervention that successfully increased rates of HCV testing among patients at risk.[22] Because existing guidelines do not concur on what factors should trigger HCV testing, we included a moderately large number of risk factors (12) in assessing the intervention. However, it was unknown which factors of the screening intervention were responsible for the screener's success and whether an abbreviated set of risk factors would be equally successful. To inform both the development of a parsimonious screening intervention and the revision of risk-based HCV testing guidelines, we examined which factors were the strongest independent predictors of testing and diagnosis of HCV.

The Hepatitis C Assessment and Testing project (HepCAT), a prospective cross-sectional

evaluation conducted in 3 urban primary care clinics, was designed to inform CDC's revision of its HCV testing recommendations. HepCAT's major goal was to evaluate an intervention designed to identify patients at risk for HCV with a PCP-implemented risk screener and test those identified as at risk. Another objective was to parse out a limited number of factors to include in a simple and effective screener. We hypothesized that using the risk screener would increase testing rates and that a brief screener incorporating fewer risk factors would perform as well as the full screener. We examined the performance of the screener overall as well as the extent to which each specific risk factor predicted HCV.

METHODS

The HepCAT project was funded by the CDC through the Agency for Healthcare Research and Quality's Accelerating Change and Transformation in Organizations and Networks (ACTION) program. ACTION promotes "field-based research designed to promote innovation in health care delivery by accelerating the diffusion of research into practice"[23] rather than funding only traditional research studies involving comparison and control groups.

The ACTION network includes 15 partnerships, each with a lead organization that serves as the prime contractor for projects. Boston University is the prime contractor for one partnership that includes the Montefiore Medical Center in the Bronx, New York, and HepCAT, a collaborative effort between these 2 organizations, was conducted in 3 clinics affiliated with the center. The 3 clinics provide approximately 150 000 primary care visits annually to more than 54 000 adults; the clinics are located in economically depressed areas of the Bronx and serve patients with high rates of poverty, substance use, and sexually transmitted diseases. The estimated prevalence of HCV infection in the adult population seen in the 3 clinics is 7.7%.[24]

Intervention

The risk-based screening intervention involved prompting physicians with a clinical reminder sticker to ask whether a patient had any of 12 specific HCV-related risks and to order HCV tests according to the presence of those risks. The sticker (Figure 1) incorporated a double-layer "carbon copy" design; once completed, the top copy of the sticker could be removed and placed in a secured box in the examination room for research purposes while the other copy remained affixed to the medical chart.

In addition to the sticker, intervention training included the following: on-site educational sessions including a standardized presentation for all PCPs and clinic staff delivered prior to and once during the intervention period, regular communication between the research team and clinical leadership, electronic provision of a weekly scientific article on HCV to all PCPs, and environmental reminders (HepCAT buttons, pocket cards, and posters). Also, project staff visited each clinic twice per week to place stickers on all progress notes, encourage adherence to screening protocols, and elicit feedback from PCPs and other clinic staff. Furthermore, each clinic had a "physician champion," a member of the research team who regularly visited the clinic to maintain PCPs' engagement in the intervention. All PCPs were supplied with a script (in English and Spanish) to standardize and normalize the introduction of the screener.

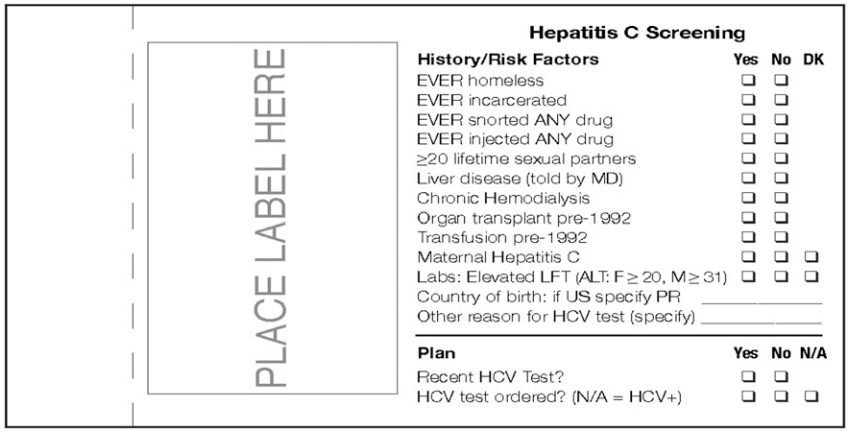
Each PCP was asked to complete the sticker at every visit for patients who had not been tested for HCV in the preceding 12 months and to order an HCV antibody test if any risk factor was identified. The risk-based screening intervention was conducted over a 15-week period from November 2008 to March 2009.

Data Collection

Given the expectation that all adults seen in the clinics would be screened, demographic and HCV testing information on each adult (18 years old or older) seen during the intervention period was extracted from the electronic medical record. Risk factor data were collected from the risk screener sticker. These data sets were merged into a single database. We compared patients who had been screened with those who had not been screened. Among screened patients, we examined risk factors, HCV testing, and HCV positivity rates.

Measures

The main outcomes in our analysis were HCV testing and HCV positivity. A patient was considered to have been tested if an anti-HCV antibody test was performed within 90 days of the clinic visit date. HCV antibody positivity was defined as a positive anti-HCV test within the same time period. To determine the effects of the screener on HCV testing and positivity, we analyzed patients who had and had not been screened separately. A patient was considered screened if a sticker was submitted with any information about HCV risk factors recorded. A patient was considered not screened if no sticker was submitted or if the submitted sticker was blank. After examining the main



Note. ALT = alanine aminotransferase; DK = don't know; HCV = hepatitis C virus; LFT = liver function test; N/A = not applicable; PR = Puerto Rico.

FIGURE 1—Clinical reminder sticker used in the risk screening intervention.

outcomes, we examined the incremental value of each element of the screener to assess its performance with respect to promoting HCV testing and identifying cases of anti-HCV positivity; analyses focused on proportions of patients tested and results for patients with each risk factor.

Data Analysis

We used Stata software in conducting our analyses.[25] We initially conducted the *t* test, the χ^2 test, the rank sum test to compare screened and nonscreened patients with respect to demographic characteristics. Among screened patients, we examined the proportions with any risk identified (vs no risk identified), the proportions who had (vs had not) been tested, and rates of HCV positivity. We then investigated the rates of testing and HCV positivity associated with each risk factor.

We identified the incremental value of each screener element through an iterative process as follows. Initially we determined which risk factor was most strongly associated with positivity and counted the number of positive cases identified by asking about that risk factor. Next, we removed the HCV-positive cases identified by the first risk factor and determined which of the other risk factors identified the most remaining positive cases. Then the HCV-positive cases identified by the second risk factor were removed and a third risk factor that identified the most remaining cases was sought. The process was repeated until the remaining risk factors did not identify any new HCV positive cases.

RESULTS

Of the 13 371 patients with at least 1 primary care visit during the intervention period, 4339 had previously been tested and 51 were missing critical data; as a result, 8981 patients were included in our analysis. Table 1 shows comparisons of patients screened and not screened during the intervention period.

The mean age of the 8981 patients was 47.8 years, and one fourth were male. More than half were Latino, about 30% were Black, and fewer than 5% were White. About half were insured via Medicaid, just under a quarter had commercial insurance, about 14% were uninsured, and 12% had Medicare. A screener was completed for 3250 (36.2%) of the patients seen. Male and White patients were more likely to be screened, and Black patients were less likely to be screened; there were no differences in screening rates according to age or insurance status.

Table 2 presents proportions of testing and yield (rate) of anti-HCV positivity among those screened and not screened and among those with and without identified risks. During the intervention, 13.1% of all patients seen in the clinics were tested for HCV. However, this level of testing was driven primarily by screening: 25.3% of screened patients were tested, as opposed to only 6.2% of unscreened patients. The yields of anti-HCV positivity were 5.9% among those with no screening documentation and 5.0% among the screened population.

Of the 3250 patients screened, 27.8% (n = 902) had at least 1 HCV risk factor. Of these patients, 55.4% (n = 500) were tested for HCV; 13.7% of tested patients had no identified risk. The yield of anti-HCV positivity among tested patients was higher for those with a risk factor (6.8%) than for those without a risk factor (2.2%).

TABLE 1—Characteristics of Patients Seen During the Intervention Period: Bronx, NY, 2008–2009

Characteristic	Total (n = 8981), No. (%)	Screened (n = 3250), No. (%)	Not Screened (n = 5731), No. (%)	*P*
Gender				.002[a]
Male	2330 (25.9)	906 (27.9)	1424 (24.8)	
Female	6651 (74.1)	2344 (72.1)	4307 (75.2)	
Race/ethnicity				.01[a]
White	389 (4.3)	272 (4.7)	117 (3.6)	.01[a]
Black	2733 (30.4)	1698 (29.6)	1035 (31.8)	.03[a]
Latino	4734 (52.7)	3025 (52.8)	1709 (52.6)	.86
Other/unknown	1125 (12.5)	736 (12.8)	389 (12.0)	.23
Insurance coverage				.09
Medicare	1029 (11.5)	645 (11.3)	384 (11.8)	.42
Medicaid	4609 (51.3)	2981 (52.0)	1628 (50.1)	.08
Commercial	2062 (23.0)	1283 (22.4)	779 (24.0)	.09
None	1272 (14.2)	819 (14.3)	453 (13.9)	.64

[a]Significant difference between groups at $P < .05$.

Identification of Risk Factors

Table 3 provides details on the 902 screened patients identified as having at least 1 HCV risk factor. The most commonly identified risk factors, documented among more than 20% of screened patients, were history of multiple sex partners, intranasal drug use, elevated ALT, homelessness, and incarceration. Histories of liver disease and blood transfusions before 1992 were reported by 10% to 15% of those screened. A smaller percentage (6.2%) of patients reported injection drug use; just over 2% reported maternal HCV. Very few patients (< 1%) reported chronic hemodialysis or organ transplants before 1992.

Testing Rates for Each Risk Factor

All prevalent risk factors predicted testing relatively well, at close to or above 50% of the time. Testing was most often conducted (> 60% of the time) among those who reported histories of multiple sex partners, transfusions, and maternal hepatitis. In addition, more than half of those with histories of incarceration, intranasal drug use, and elevated ALT were tested, as were almost half of those with histories of homelessness, injection drug use, and liver disease.

Yield of Patients Tested

Overall, the yields of HCV testing were high. Sixty-three percent of patients with a history of injection drug use had positive anti-HCV test

TABLE 2—Hepatitis C Virus (HCV) Screening, Testing, and Yield During the Intervention Period: Bronx, NY, 2008–2009

	Tested for HCV, No. (%)	Yield (Anti-HCV Positivity), No. (%)
Total seen (n = 8981)	1179 (13.1)	62 (5.3)
Not screened (n = 5731)	357 (6.2)	21 (5.9)
Screened (n = 3250)	822 (25.3)	41 (5.0)
No risk (n = 2348)	322 (13.7)	7 (2.2)
Any risk (n = 902)	500 (55.4)	34 (6.8)

results. Positivity was also high among those reporting liver disease (26.8%), maternal HCV (16.7%), intranasal drug use (15.7%), incarceration (10.9%), homelessness (9.0%), multiple sex partners (8.4%), elevated ALT (7.3%), and transfusions (6.9%).

Incremental Value of Screener Elements

Table 4 illustrates the incremental predictive value of each risk factor starting with the factor with the highest yield: injection drug use. Injection drug use was the strongest predictor of anti-HCV positivity; sole inclusion of the "ever injected drugs" variable would have predicted 41.5% of identified cases. Intranasal drug use was the second strongest predictor; the incremental benefit of adding "ever snorted drugs" as a second factor would have led to the additional identification of 14.6% of cases, for a total yield from a 2-question screener of 56.1% of the identified cases of anti-HCV positivity. Ultimately, we determined that a 7-element screener that comprised injection drug use, intranasal drug use, elevated ALT, transfusions before 1992, maternal HCV, 20 or more lifetime sex partners, and existing liver disease would have accounted for all anti-HCV cases identified.

DISCUSSION

We found that anti-HCV testing increased in a primary care setting when a 12-item risk screener was implemented and that 10 elements identified patients at risk. However, 6 of the elements would have identified the same number of cases of anti-HCV positivity, suggesting that a briefer screening instrument would be equally effective.

Our findings demonstrate the utility of risk-based testing in identifying individuals positive for anti-HCV, confirming the findings of previous studies.[20,21] Earlier research showed that the 3 study clinics had already tested 39.7% of their patient populations and had estimated that 59.7% of these patients were anti-HCV positive before the risk screening intervention.[24] Although it could be argued that the individuals identified before the intervention would have presented with the most obvious risks, the risk screener still identified an additional 13.1% for screening in the previously untested group, and of these patients 62 were anti-HCV positive. This increase in testing was primarily because of the screened population; 25.3% of screened patients were subsequently tested for HCV, whereas the testing rate among unscreened patients was 6.2%. More than half of patients who had at least 1 risk factor identified on the screener were tested. Thus, the intervention was effective in increasing testing rates both overall and, particularly, among patients with identified risk factors.

It is notable that the testing rate among patients identified as having HCV risk factors was not closer to 100%, particularly in the case of risk factors such as injection drug use that are widely known to be associated with HCV. There may be several reasons for this finding, including the steps required by both patient and PCP to complete the anti-HCV testing process. For patients, testing required going to a separate area in the clinic and waiting to have blood drawn; for PCPs, ordering tests required completing a lab slip. These may be important barriers to be mindful of in attempts to increase testing in primary care settings,

TABLE 3—Risk Factor Identification, Hepatitis C Virus (HCV) Testing, and Yield: Bronx, NY, 2008–2009

Factor	Risk Factor Identified, No. (%)	Tested for HCV, No. (%)	Yield (Anti-HCV Positivity), No. (%)
Ever homeless	234 (25.9)	111 (47.2)	10 (9.0)
Ever incarcerated	214 (23.7)	119 (55.6)	13 (10.9)
Ever snorted drugs	252 (27.9)	134 (53.2)	21 (15.7)
Ever injected drugs	56 (6.2)	27 (48.2)	17 (63.0)
≥ 20 lifetime sex partners	270 (29.9)	167 (61.9)	14 (8.4)
Liver disease (physician diagnosis)	115 (12.7)	56 (48.7)	15 (26.8)
Chronic hemodialysis	9 (1.0)	2 (22.2)	0 (0.0)
Transplant before 1992	5 (0.5)	2 (40.0)	0 (0.0)
Transfusion before 1992	108 (12.0)	72 (66.7)	5 (6.9)
Maternal hepatitis C	19 (2.1)	12 (63.2)	2 (16.7)
Elevated alanine aminotransferase (documented in electronic medical record)	242 (26.8)	137 (56.6)	10 (7.3)

Note. Totals sum to more than 902 because patients often had multiple factors identified. The sample size was n = 902.

TABLE 4—Incremental Value of Hepatitis C Virus (HCV) Screening Items: Bronx, NY, 2008–2009

Factor	No. of Patients With Identified HCV Risk	No. of Patients With Positive Test Results (% of Positive Cases Overall)	Cumulative %
Ever injected drugs	56	17 (41.5)	41.5
Ever snorted drugs	200	6 (14.6)	56.1
Elevated alanine aminotransferase (documented in electronic medical record)	185	4 (9.8)	65.9
Transfusion before 1992	59	3 (8.0)	73.1
≥ 20 lifetime sex partners	115	2 (4.9)	78.0
Maternal hepatitis C	10	1 (2.4)	80.5
Liver disease (physician diagnosis)	23	1 (2.4)	82.9
Ever homeless	66	0 (0.0)	82.9
Ever incarcerated	67	0 (0.0)	82.9
Chronic hemodialysis	0	0 (0.0)	82.9
Transplant before 1992	0	0 (0.0)	82.9
Total		34/41	82.9

Note. The cumulative percentage does not reach 100% because 7 of the 41 patients with positive anti-HCV results had no risk factors identified on the risk screener.

particularly given the myriad screening and preventive care activities that are already recommended.

In addition, there was no requirement in the HepCAT protocol for patients to be tested at the visit when the sticker was placed in the chart, and the PCP may have planned to test the patient at a later visit. Only the Veterans Health Administration, via a federal mandate and real-time electronic clinical reminders, has developed a mechanism to fully implement risk-based HCV screening and testing in primary care settings[23]; an electronic reminder to screen might be the only method to achieve full implementation.

A key question in the development and implementation of the HepCAT intervention was which items to include in the screener. Although screened patients reported high rates of intranasal and injection drug use, as well as multiple sex partners, homelessness, incarceration, liver disease, and transfusions, our analysis illustrated that a screener with fewer risk factors could be as effective as the 12-item screener. Our results showed that two thirds of patients with positive anti-HCV test results were identified with a screener that included just 3 factors (injection drug use, intranasal drug use, elevated ALT); 4 additional factors (transfusions, maternal hepatitis C, 20 or more lifetime sex partners, liver disease) identified an additional 17% of cases.

It appears that some factors, such as incarceration and homelessness, may actually be proxies for the other risk factors and will not produce additional benefits in terms of identifying cases of anti-HCV positivity. However, because 17% of the patients with positive anti-HCV test results had no risk factors identified on the screener, it will be important to understand the characteristics of patients without risk factors who were tested and the reasons they were tested, including assessing how demographic differences may or may not relate to other risk factor differences. For example, it is important to examine reasons why male patients were more likely to be screened than were female patients and White patients were more likely to be screened than were Black patients.

Our findings mirror recent work showing that a parsimonious screener can be effective[1,21] and practical in the context of routine care. Consistent with our results, Zuniga et al.,[21] using retrospective data on veterans, found that screening only for injection drug use would have identified 41% of cases of anti-HCV positivity in that population; in addition, they found that a risk screener including only 5 factors (injection drug use, blood transfusion before 1992, service during the Vietnam era, tattoos, and history of abnormal liver function tests) and a risk screener incorporating the 5 factors independently associated with anti-HCV positivity would have identified 97% of cases with 20% fewer individuals being tested. Our results also are consistent with an analysis of National Health and Nutrition Examination Study data conducted by Armstrong et al.,[1] who found that injection drug use, elevated ALT, and transfusions before 1992 identified 85% of cases of anti-HCV positivity.

An important difference between our work and previous studies is our inclusion of intranasal drug use as a screening item. We found that although intranasal drug use was not independently associated with anti-HCV positivity (probably as a result of sample size limitations), it identified almost 15% of cases of positivity. Given our findings and the possibility that patients will be more likely to acknowledge intranasal than injection drug use owing to the stigma often associated with injection drug use, we propose that intranasal drug use be considered for inclusion in brief screeners.

Overall, it is important to note that our study, conducted in the context of routine primary care in a high-risk population, is congruent with the results obtained by Armstrong et al. in their US population-based sample, and it appears clear that a brief screener including injection drug use, elevated ALT, and transfusions before 1992 will be effective in identifying HCV

with or without the inclusion of intranasal drug use. Our results also concur with research by McGinn et al.[20] supporting risk-based testing; however, their study involved a 27-item questionnaire, whereas our focus on using a brief screener is more practical for implementation.

With the current policy focus on the medical home model and the role of PCPs in coordinating and taking responsibility for all aspects of a patient's care, the expectation will remain for PCPs to do more with less time. Thus, the challenge of integrating a new intervention to identify HCV in primary care settings remains. To minimize the impact on already-overtaxed PCPs, it is crucial to identify those elements most predictive of HCV positivity. We found that PCPs can perform effective HCV screening with a screener that includes many fewer risk factors than previously reported in the literature. Moreover, screening could be performed by ancillary providers, such as nursing staff, and this type of intervention could be easily generalized to other settings or other clinical conditions given its relative simplicity and the lack of technology required for implementation.

Limitations

Our study involved some limitations. First, it was difficult to sustain a high level of PCP adherence to the intervention, and PCPs were unable to do more than check a box on a screener (e.g., identify country of birth if outside the United States). Thus, it is important to consider whether screening could be conducted by other clinical staff. Second, because we were unable to track unused laboratory slips, we cannot determine whether untested patients with risk factors were referred for but did not undergo testing or whether these patients were not referred.

Third, this study was observational; without a comparison group, we cannot establish a causal link between the intervention and the increase in testing. Finally, the small number of cases of identified anti-HCV positivity limited our examination of the incremental value of each of the risk factors in creating a briefer risk screener.

Conclusions

We found that a brief risk screener with a paper-based clinical reminder was effective in increasing HCV testing in a primary care setting. With more effective treatments now available, it is critical that the process of identification of HCV be improved, given that care and treatment cannot be offered without diagnosis. Primary care is the front line of health care for most patients and the optimal location for simple risk screening.

Given the many challenges facing PCPs and the numerous preventive care activities expected of them, future efforts should focus on testing a more parsimonious risk screener and determining whether ancillary staff could conduct screening activities. In addition, there is a need for future research testing the use of HCV screening by different types of PCPs, including studies involving experimental designs. Studies are also needed to explore the actual utility of a brief screener, followed by validation of that screener. Finally, we recommend that the cost-effectiveness of our intervention be assessed; if it is cost-effective for HCV screening, it might serve as a model for primary care screening of other undiagnosed clinical conditions. ■

About the Authors

Mari-Lynn Drainoni, Elisa A. Koppelman, and Cindy L. Christiansen are with the Department of Health and Policy Management, Boston University School of Public Health, Boston, MA. Alain H. Litwin is with the Division of General Internal Medicine, Department of Medicine, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY. Bryce D. Smith and Cindy M. Weinbaum are with the Division of Viral Hepatitis, National Center for HIV/Viral Hepatitis/STD/TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA. M. Diane McKee is with the Department of Family Medicine, Albert Einstein College of Medicine and Montefiore Medical Center. Allen L. Gifford is with the VA HIV/Hepatitis Quality Enhancement Research Initiative, Edith Nourse Rogers Memorial Veterans Administration Hospital, Bedford, MA. William N. Southern is with the Division of Hospital Medicine, Department of Medicine, Albert Einstein College of Medicine and Montefiore Medical Center.

Correspondence should be sent to Mari-Lynn Drainoni, PhD, 715 Albany St, T3W, Boston, MA 02118 (e-mail: drainoni@bu.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This article was accepted December 22, 2011.

Contributors

M.-L. Drainoni designed the study, led the writing, and identified analyses to be conducted. A. H. Litwin co-designed the study and assisted in data analysis and interpretation. B. D. Smith assisted in data interpretation and participated in the writing of the article. E. A. Koppelman coordinated all study activities and participated in the writing and editing of the article. M. D. McKee led the engagement of the clinics and contributed to the editing of the article. C. L. Christiansen served as the study biostatistician, assisted in data interpretation, and contributed to the editing of the article. A. L. Gifford served as a clinical resource for the study and assisted in the study design. C. M. Weinbaum conceptualized the study and contributed to the editing of the article. W. N. Southern led the study analysis and data interpretation and played a leadership role in editing the article.

Acknowledgments

This study was funded by the Agency for Health Care Research and Quality (AHRQ) and the Centers for Disease Control and Prevention via the AHRQ Accelerating Change and Transformation in Organizations and Networks (ACTION) initiative (contract HHSA2902006000012 T0#4 to Boston University).

Note. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the AHRQ.

Human Participant Protection

This study was approved by the institutional review boards of the Centers for Disease Control and Prevention, the Boston University Medical Center, and the Montefiore Medical Center. A waiver of informed consent was granted for all study activities.

References

1. Armstrong GL, Wasley A, Simard EP, McQuillan GM, Kuhnert WK, Alter MJ. The prevalence of hepatitis C virus infection in the United States, 1999 through 2002. *Ann Intern Med.* 2006;144(10):705–714.

2. Kanwal F, Schnitzler MS, Bacon BR, Hoang T, Buchanan PM, Asch SM. Quality of care in patients with chronic hepatitis C virus infection. *Ann Intern Med.* 2010;153(4):231–239.

3. Ly K, Xing J, Klevens M, Jiles R, Ward J, Holmberg S. The growing burden of mortality from viral hepatitis in the US, 1999–2007. Paper presented at: annual meeting of the American Association for the Study of Liver Diseases, November 2011, San Francisco, CA.

4. Volk ML, Tocco R, Saini S, Lok AS. Public health impact of antiviral therapy for hepatitis C in the United States. *Hepatology.* 2009;50(6):1750–1755.

5. Kwo PY, Lawitz EJ, McCone J, et al. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naïve patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. *Lancet.* 2010;376(9742):705–716.

6. Roblin DW, Smith BD, Weinbaum CM, Sabin M. Hepatitis C virus screening practices and prevalence in an MCO, 2000–2007. *Am J Manag Care.* 2011;17(8):548–555.

7. Wasley A, Finelli L, Bell BP, Alter MJ. The knowledge and behavior of HCV-infected persons identified in a national seroprevalence survey, United States, 2001–2004. Paper presented at: 12th International Symposium on Viral Hepatitis and Liver Disease, July 2006, Paris, France.

8. Centers for Disease Control and Prevention. Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. *MMWR Recomm Rep.* 1998;47(RR19):1–39.

9. Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology.* 2009;49(4):1335–1374.

10. Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C. *Hepatology.* 2004;39(4):1147–1171.

11. US Dept of Veterans Affairs. Hepatitis C testing and prevention counseling guidelines for VA health care practitioners. Available at: http://www.hepatitis.va.gov/provider/guidelines/testing-prevention-counseling.asp. Accessed July 24, 2012.

12. Dienstag JL, McHutchinson JG. American Gastroenterological Association medical position statement on the management of hepatitis C. *Gastroenterology.* 2006;130(1):225–230.

13. National Institutes of Health. National Institutes of Health Consensus Development Conference statement: management of hepatitis C: 2002. *Hepatology.* 2002; 36(suppl 1):S3–S20.

14. Ko YC, Ho MS, Chiang TA, Chang SJ, Chang PY. Tattooing as a risk of hepatitis C virus infection. *J Med Virol.* 1992;38(4):288–291.

15. Tumminelli F, Marcellin P, Rizzo S, et al. Shaving as potential source of hepatitis C virus infection. *Lancet.* 1995;345(8950):658.

16. Hayes MO, Harkness GA. Body piercing as a risk factor for viral hepatitis: an integrative research review. *Am J Infect Control.* 2001;29(4):271–274.

17. Conry-Cantilena C, VanRaden M, Gibble J, et al. Routes of infection, viremia and liver disease in blood donors found to have hepatitis C virus infection. *N Engl J Med.* 1996;334(26):1691–1696.

18. Tortu S, Neaigus A, McMahon J, Hagen D. Hepatitis C among non-injecting drug users: a report. *Subst Use Misuse.* 2001;36(4):523–534.

19. Zalumas J, Rose C. Hepatitis C and HIV in incarcerated populations: fights, bites, searches, and syringes! *J Assoc Nurses AIDS Care.* 2003;14(suppl 5): 108S–115S.

20. McGinn T, O'Connor-Moore N, Alfandre D, Gardenier D, Wisnivesky J. Validation of a hepatitis C screening tool in primary care. *Arch Intern Med.* 2008;168(18):2009–2013.

21. Zuniga IA, Chen CC, Lane DS, Allmer J, Jimenez-Lucho VE. Analysis of a hepatitis C screening programme for US veterans. *Epidemiol Infect.* 2006;134(2): 249–257.

22. Litwin AH, Smith BD, Drainoni ML, et al. Primary care-based interventions are associated with increases in hepatitis c virus testing for patients at risk. *Dig Liver Dis.* 2012;44(6):497–503.

23. Agency for Health Care Research and Quality. Accelerating Change and Transformation in Organizations and Networks (ACTION): field partnerships for applied research. Available at: http://www.ahrq.gov/research/action.htm. Accessed July 24, 2012.

24. Southern WN, Drainoni M, Smith BD, et al. Hepatitis C testing practices and prevalence in a high-risk urban ambulatory care setting. *J Viral Hepat.* 2011;18(7): 474–481.

25. *Stata, Version 10.0.* College Station, TX: StataCorp LP; 2007.

BMJ

BMJ 2012;344:e1442 doi: 10.1136/bmj.e1442 (Published 12 March 2012)

RESEARCH

Inverse association between cancer and Alzheimer's disease: results from the Framingham Heart Study

Jane A Driver *assistant professor of medicine*[1 2 3], Alexa Beiser *professor of neurology and biostatistics*[4 5 6], Rhoda Au *associate professor of neurology*[4], Bernard E Kreger *assistant professor of medicine*[6 7], Greta Lee Splansky *director of operations, Framingham Heart Study*[6], Tobias Kurth *director of research*[3 8], Douglas P Kiel *professor of medicine*[9], Kun Ping Lu *professor of medicine*[9], Sudha Seshadri *associate professor of neurology*[4], Phillip A Wolf *professor of neurology*[4 6]

[1]Geriatric Research Education and Clinical Center and Boston VA Medical Center, Boston, MA 02130, USA; [2]Division of Aging, Brigham and Women's Hospital, Boston, MA, USA; [3]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA; [4]Department of Neurology, School of Medicine, Boston University, Boston, MA, USA; [5]Department of Biostatistics, School of Public Health, Boston University, Boston, MA, USA; [6]Framingham Heart Study, National Heart, Lung, and Blood Institute, Framingham, MA, USA; [7]Department of Medicine, Boston University, Boston, MA, USA; [8]INSERM Unit 708—Neuroepidemiology, Bordeaux, France, and University of Bordeaux, France; [9]Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

Abstract

Objectives To relate cancer since entry into the Framingham Heart Study with the risk of incident Alzheimer's disease and to estimate the risk of incident cancer among participants with and without Alzheimer's disease.

Design Community based prospective cohort study; nested age and sex matched case-control study.

Setting Framingham Heart Study, USA.

Participants 1278 participants with and without a history of cancer who were aged 65 or more and free of dementia at baseline (1986-90).

Main outcome measures Hazard ratios and 95% confidence intervals for the risks of Alzheimer's disease and cancer.

Results Over a mean follow-up of 10 years, 221 cases of probable Alzheimer's disease were diagnosed. Cancer survivors had a lower risk of probable Alzheimer's disease (hazard ratio 0.67, 95% confidence interval 0.47 to 0.97), adjusted for age, sex, and smoking. The risk was lower among survivors of smoking related cancers (0.26, 0.08 to 0.82) than among survivors of non-smoking related cancers (0.82, 0.57 to 1.19). In contrast with their decreased risk of Alzheimer's disease, survivors of smoking related cancer had a substantially increased risk of stroke (2.18, 1.29 to 3.68). In the nested case-control analysis, participants with probable Alzheimer's disease had a lower risk of subsequent cancer (0.39, 0.26 to 0.58) than reference participants, as did participants with any Alzheimer's disease (0.38) and any dementia (0.44).

Conclusions Cancer survivors had a lower risk of Alzheimer's disease than those without cancer, and patients with Alzheimer's disease had a lower risk of incident cancer. The risk of Alzheimer's disease was lowest in survivors of smoking related cancers, and was not primarily explained by survival bias. This pattern for cancer is similar to that seen in Parkinson's disease and suggests an inverse association between cancer and neurodegeneration.

Introduction

Limited data suggest that cancer survivors have a decreased risk of Alzheimer's disease and that people with Alzheimer's disease have lower rates of cancer.[1-6] Evidence of an inverse relation between Parkinson's disease and most cancers is now convincing.[7-13] A link between cancer and neurodegeneration is plausible as they share several genes and biological pathways, including inappropriate activation and deregulation of the cell cycle.[14-22] Signaling along these pathways results in opposite end points: in the case of cancer, uncontrolled cell proliferation, and in the case of neurodegeneration, apoptotic cell death. Proteins such as p53, a major regulator of apoptosis, and Pin1, which has a dual role in cell cycle control and protein folding, play a key part in the pathophysiology of both Alzheimer's disease and cancer.[15] A better understanding of the biological links between these two families of diseases is already opening new therapeutic horizons.

In one population based cohort study, people with prevalent cancer had a 43% lower risk of ever developing Alzheimer's disease, and those with prevalent Alzheimer's disease had a 69% lower risk of being admitted to hospital for cancer.[4] Although these results are intriguing, establishing a relation

Correspondence to: J A Driver jdriver@partners.org

between age related diseases is complex, and several issues must be dealt with before concluding that this represents a true association.[23] Because severe cognitive impairment leads to decreased screening and reporting of cancer symptoms,[24 25] it is difficult to know to what extent lower cancer rates in people with Alzheimer's disease are caused by decreased incidence or under-diagnosis. For this reason, assessing the incidence of Alzheimer's disease among cancer survivors who are cognitively intact at baseline is the preferable analysis. Here the major challenge is the problem of selective mortality, since cancer survivors may have a lower risk of Alzheimer's disease simply because they are more likely to die before they can develop it. Available studies have not sufficiently tackled the problems of selective mortality. Additional limitations of previous analyses are the exclusion from analysis of incident cancers that develop after baseline but before Alzheimer's disease and reliance on self report or medical record systems to identify cancer. We investigated the relation between cancer and Alzheimer's disease using data from the Framingham Heart Study, a prospective cohort with frequent examinations of participants, prospective validation of both Alzheimer's disease and cancer, and over 50 years of follow-up.

Methods

The Framingham Heart Study is a longitudinal community based cohort study of cardiovascular risk factors that started in 1948 in Framingham, Massachusetts, United States. The original cohort comprised 5209 participants (2336 men and 2873 women) aged 28-62 at the first examination. In 1971, children of the original cohort and their spouses were recruited to form the offspring cohort (5214 participants). Participants have undergone direct evaluations, including a medical history, physical examination, and laboratory testing every two years in the original cohort and about every four years in the offspring cohort. The study design and entry criteria for both cohorts have been described in detail elsewhere.[26] Participants of the original cohort aged 65 and older who were free of dementia and attended examination cycle 20 (1986-90) comprise the sample for our primary investigation (n=1278). We followed these participants for incident dementia for a mean of 10 years. The nested case-control study comprised participants from both the original and the offspring cohorts (n=1485). All participants gave written informed consent.

Ascertainment of cancer cases

We identified possible cancer cases in the Framingham study at routine examinations or, if participants did not attend an examination, by postal surveys or telephone interviews for updates on health history. Cases were also identified through surveillance of admissions to the local Framingham hospital and from death records. Once a case was identified, we confirmed the diagnosis from the patient's medical records, including pathology reports. Two independent people reviewed the medical records. Most cancers were confirmed by pathology reports, and fewer than 3.4% of diagnoses were based solely on death certificates or clinical diagnoses.[27] We coded primary cancers using the World Health Organization ICD-O (international classification of diseases) classification. For this analysis we did not include non-malignant neoplasms and non-melanoma skin cancers in the definition of cancer.

Ascertainment of dementia cases

Since examination cycle 14 in the original cohort and examination 2 in the offspring cohort, a dementia-free cohort of 7809 participants has been under continuous surveillance for the development of dementia and Alzheimer's disease. The Folstein mini-mental state examination[28] was administered at regular cycle examinations, and participants who scored below an education based cut-off point or had a 3 point decrement in their score from a preceding examination (or 5 point decrement overall) were referred for more indepth testing. Participants also underwent an indepth evaluation if they or someone in their family reported symptoms of memory loss or were referred by a Framingham study physician or staff member for evaluation of neurological symptoms. A panel of at least one neurologist and one neuropsychologist determined cases of dementia, dates of diagnosis, and subtypes using data, where available, from the neurologist's examination, neuropsychological test performance, Framingham study records, hospital records, information from primary care physicians, interviews with the families, computed tomography and magnetic resonance imaging records, and confirmation of autopsy findings. Dementia was required to be present for at least six months of follow-up before the diagnosis was confirmed. All participants identified as having dementia had at least mild severity by the clinical dementia rating score of 1 or more. Cases of Alzheimer's disease met the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association criteria for possible and probable Alzheimer's disease.[29] We classified dementia as any dementia (met the criteria for any dementia, including Alzheimer's disease), possible Alzheimer's disease (met the clinical criteria for Alzheimer's disease but with an atypical course or evidence of a second process contributing to the dementia), and probable Alzheimer's disease (met the criteria for Alzheimer's disease without evidence of another process contributing to the dementia).

Assessment of covariates

From the baseline visit (examination 20) we obtained data on personal characteristics (age, sex, and education) and laboratory test results (homocysteine levels and apolipoprotein E genotype). We collected information on cancer risk factors (tobacco use and body mass index) at baseline and updated this periodically throughout the study. For the case-control study, we used the covariates closest to the matching date.

Statistical analyses

Prospective cohort study

Participants contributed up to 22 years of follow-up from the baseline examination to the development of dementia, death, or the last evaluation. We used Cox proportional hazard models to determine the hazard ratio and 95% confidence intervals for the risk of any dementia, any Alzheimer's disease (possible or probable), and probable Alzheimer's disease in those with and without a history of verified cancer. The cancer history variable was updated to include cases of incident cancer that occurred during follow-up after baseline. We carried out separate analyses for participants with any cancer (excluding non-melanoma skin cancer), smoking related cancers (oral, pharynx, larynx, oesophagus, stomach, pancreas, lung, cervix, bladder, and kidney)[30], and non-smoking related cancers. The primary models were adjusted for age, sex, and smoking. We then repeated the analysis in a smaller subset of patients with data on other risk factors for Alzheimer's disease: apolipoprotein E4 status, educational level, and plasma homocysteine level.[30-32]

To explore whether the relation between cancer and Alzheimer's disease might be mainly due to selective mortality, we first

restricted the analysis to participants who survived at least to age 80. If the association was predominantly due to the death of cancer survivors, then it should be diminished when these patients were excluded. We then investigated the relation between history of cancer at examination 20 (baseline) and a different neurological outcome—namely, the subsequent risk of incident stroke. If a decreased risk of cancer in patients with Alzheimer's disease was primarily due to their increased mortality rate, then they would be expected to have a lower rate of stroke as well.

Nested case-control study

To evaluate the relation between dementia and subsequent cancer, we matched each dementia case with up to three controls of the same age and sex who were free of dementia at the time of dementia diagnosis of the case (index date). Both cases and reference participants were free of cancer as of the index date. Participants who eventually developed dementia were considered potential controls up to five years before their data of dementia diagnosis (to avoid subclinical dementia in controls). We used Cox models to determine the hazard of incident cancer in dementia cases compared with the reference participants. Models were adjusted for tobacco use and body mass index. Data on covariates was taken from the examination closest to the index date. We calculated the cumulative incidence of cancer in those with and without Alzheimer's disease adjusted for age, sex, body mass index, and smoking.

Results

At examination 20 (baseline), 1278 participants (38.8% men) aged 65 or older had normal cognitive status. The mean age among cancer survivors (n=176) was 77 years and among those with no history of cancer (n=1102) was 76 years (table 1⇓). Educational level, positivity for apolipoprotein E4, and homocysteine level did not differ substantially between those with and without cancer. Overall, 323 cases of dementia were diagnosed over a mean of 10 years of follow-up. Of these, 221 (86%) met the criteria for probable Alzheimer's disease and 36 for possible Alzheimer's disease (18 had Alzheimer's disease with stroke and 18 had both Alzheimer's disease and vascular dementia). Of the 66 patients with non-Alzheimer's disease, 24 had Lewy body dementia, 15 had vascular dementia, two had frontotemporal dementia, and 25 were classified as having "other" dementia.

Cancer history and risk of Alzheimer's disease

At baseline 176 participants had a history of cancer, and during follow-up an additional 247 people were diagnosed as having incident cancer but before the diagnosis of Alzheimer's disease. Cancers at baseline were more likely to be screening related and non-smoking related than incident cancers (table 2⇓). Cancers with high lethality such as those of the lung, pancreas, and brain were much more common among incident cancers. Cancer survivors had a substantially lower risk of probable Alzheimer's disease (hazard ratio 0.67, 95% confidence interval 0.47 to 0.97), adjusted for age, sex, and smoking (table 3⇓), but the lower risks of any Alzheimer's disease (0.81) and any dementia (0.83) did not reach statistical significance. The risk of probable Alzheimer's disease was lower among survivors of smoking related cancers (0.26, 0.08 to 0.82) than among those with non-smoking related cancers (0.82, 0.57 to 1.19). Further adjustment for education, apolipoprotein E4 genotype, and homocysteine level gave slightly higher risk estimates, but the loss of statistical significance was possibly due to limited power in this smaller subset (table 3). Limiting the analysis to participants who survived at least to age 80 did not change the inverse association between cancer history and probable Alzheimer's disease (0.68, 0.47 to 0.99, table 4⇓). In contrast with their substantially lower risk of Alzheimer's disease, survivors of smoking related cancer had an increased risk of incident stroke (2.18, 1.29 to 3.68), which was also seen among those who survived to age 80 (2.25, 1.29 to 3.95).

Incident Alzheimer's disease and risk of cancer

Overall, 495 cases of any dementia, 49 of possible Alzheimer's disease, and 327 of probable Alzheimer's disease were prospectively identified. Up to three dementia-free controls were matched to each case. Overall, 41 cases (8%) and 211 controls (14%) developed incident cancer. Table 5⇓ shows the frequency of individual cancer types by case-control status. In total, 6.9% (n=102) of controls were diagnosed as having screening related cancer compared with 3.4% (n=17) of cases. In the age matched group, the risk of subsequent cancer was substantially decreased for those with any dementia (0.44, 0.32 to 0.61), any Alzheimer's disease (0.45, 0.24 to 0.84), and probable Alzheimer's disease (0.39, 0.26 to 0.58, table 6⇓). When adjusted for smoking and body mass index, the decreased risk of cancer remained statistically significant and became even lower. Over the follow-up period, 57 controls (17%) developed probable Alzheimer's disease.

Discussion

In this prospective cohort study, cancer survivors had a 33% decreased risk of developing probable Alzheimer's disease compared with people without cancer. The "protective effect" of previous cancer was greater for smoking related cancers than for non-smoking related cancers. The inverse association did not change when participants who died were excluded, and was not seen when stroke was used as an alternative outcome, suggesting it was not simply an artefact of decreased survival in patients with cancer. In the case-control analysis, patients with probable Alzheimer's disease had a 61% decreased risk of incident cancer. When all dementias were included, the risk became slightly higher. Patients with dementia were less likely to develop screening related cancers than those without dementia, suggesting that at least some of the decreased risk is because of under-diagnosis.

Strengths and limitations of the study

Our analysis has several important strengths. We prospectively defined cancer and Alzheimer's disease in the Framingham Heart Study over its more than 50 years of follow-up, so that our cases represent incident disease. Nearly all malignancies in the population were captured. We updated our cancer variable to include the many incident cases that occurred after baseline, during follow-up. This strengthened the negative association between cancer and Alzheimer's disease (data not shown), and suggests that analyses that do not include incident cancer may underestimate the association. Finally, we carried out two specific analyses to tackle the critical problem of survival bias.

Several limitations should also be considered. Firstly, our cohort, by definition, did not include people who died of cancer before study baseline. It is unknown whether people who died of cancer before the age of 65 would have had the same risk as those of survivors. As in any longitudinal study, participants with long follow-up represent a select population, and this should be kept

in mind when interpreting our results. Even complex techniques used for missing data are not able to fully correct for this selection bias.[33] However, the inverse association observed in this study was not primarily due to survival bias, at least during follow-up. We did not have the power to look at the relation between Alzheimer's disease and individual cancer types, or to compare the relation between cancer and other types of dementia. Cohorts in which both Alzheimer's disease and cancer are carefully defined are by definition not of the size required for such analyses. We were unable to stratify our results by race, as the participants in our study were mainly white. Some cases of vascular dementia were present in those who met the criteria for possible Alzheimer's disease, but, if anything, this would have biased our results towards the null. Finally, owing to the relatively small numbers of patients with cancer in our analysis, we were not able to stratify analyses by whether patients had or had not received treatment for cancer.

Comparison with previous studies

Our findings are similar to those of the only two available prospective studies. In a longitudinal memory cohort of 594 patients, those with prevalent cancer at baseline seemed to develop Alzheimer's disease at a lower rate (hazard ratio 0.40, 95% confidence interval 0.12 to 1.13) than those with no history of cancer.[3] In an analysis of the cognition cohort of the Cardiovascular Health Study (n=3020), a history of cancer at baseline conferred a lower risk of probable Alzheimer's disease (hazard ratio 0.57, 95% confidence interval 0.36 to 0.90).[4] Similar to their results, we found that the inverse relation became stronger as we excluded non-Alzheimer's and mixed dementias. Our risk estimates for incidence of cancer among those with and without Alzheimer's disease were essentially the same as that of Roe and colleague's study (0.39, 0.21 to 0.74).[3] The inverse relation decreased only slightly when we included all other dementias, and it was clear in our cohort that patients with dementia had fewer screening related cancers. This suggests that under-diagnosis in people with cognitive impairment explains at least part of the decreased incidence of cancer in patients with Alzheimer's disease. However, it is unlikely to explain all of it. Two studies found a substantially decreased prevalence of cancer on autopsy among patients with Alzheimer's disease compared with age matched controls without dementia.[1 5] In their second study, Roe and colleagues found that admission to hospital for cancer was low among patients with Alzheimer's disease (hazard ratio 0.31, 0.12 to 0.86) but not in those with vascular dementia. Finally, in a unique study of 2222 Japanese survivors of the atomic bomb, people with clinically diagnosed Alzheimer's disease had a 70% decreased risk of previous cancer compared with age matched participants without dementia, whereas those with vascular dementia had a fourfold increased risk of cancer. Clearly it will never be possible to completely adjust for reporting and diagnostic bias in elderly patients with dementia, but together these findings suggest a specific inverse relation between cancer and Alzheimer's disease.

We found that survivors of smoking related cancer had a substantially lower risk of Alzheimer's disease than survivors of non-smoking related cancer—a pattern similar to that seen in Parkinson's disease. In a recent meta-analysis of 107 598 patients with Parkinson's disease, the aggregate relative risk was 0.73 (95% confidence interval 0.63 to 0.83) for any cancer, 0.61 (0.58 to 0.65) for smoking related cancers, and 0.80 (0.77 to 0.84) for non-smoking related cancers.[7] A lower rate of smoking related cancers might be expected in patients with Parkinson's disease owing to its well established negative association with smoking,[34] but, in the case of Alzheimer's disease, the association with smoking was strongly positive.[35] In our analysis, higher mortality among those with smoking related cancer did not explain the lower risk of Alzheimer's disease. Perhaps characteristics that allow someone to survive a smoking related cancer are particularly protective against neurodegeneration. That Alzheimer's disease and Parkinson's disease are associated with the same unusual pattern for cancer suggests the presence of as yet uncovered biological connections between neurodegeneration and carcinogenesis.

Biological plausibility

Although cancer survivors may have some protection from neurodegenerative diseases, the more hopeful significance of these findings will come from the underlying biology that explains this relation.[14 16] A genetic propensity against apoptosis might protect people from cancer while increasing their risk of neurodegeneration, as seen in some polymorphisms of the tumour suppressor gene p53.[36] A specific link between Alzheimer's disease and cancer is the protein Pin1, a unique enzyme that plays a part in protein folding as well as cell cycle control.[15 17 37 38] Many tumours in humans over-express Pin1,[39] whereas its function is low in the brain tissue of people with Alzheimer's disease.[39-42] As Pin1 is necessary for cell division, its inhibition causes regression of tumours,[43] whereas in mouse models of Alzheimer's disease its upregulation in postnatal neurons reverses neurodegeneration.[44] Moreover, Pin1 promoter single nucleotide polymorphisms that inhibit Pin1 expression are associated with an increased risk of Alzheimer's disease[45] but a decreased risk of cancer.[46] Drugs that can modulate Pin1 are being sought as novel therapeutic agents.

An interesting unanswered question is whether treatment for cancer modulates the risk of Alzheimer's disease. In theory, chemotherapy might protect neurons susceptible to Alzheimer's disease by suppressing inflammation[47] or blocking entry into the cell cycle,[48] both key steps in the pathway of neurodegeneration. However, to our knowledge there is no epidemiological evidence linking chemotherapy to a decreased risk of Alzheimer's disease. On the other hand, cognitive impairment is a well described complication of chemotherapy and has been associated with long term changes in brain structure and function in animal models.[49] Survivors of breast cancer who received adjuvant chemotherapy were more likely to develop dementia in one study[50] but not in another.[51] Cranial irradiation is clearly associated with neuronal damage and loss, not neuroprotection.[52] Thus, although we did not account for cancer therapy in our analysis, available evidence suggests that, if anything, it would have biased our results towards the null.

Conclusions and future work

The results of our analysis support the possibility of a true inverse relation between cancer and Alzheimer's disease. This study is, however, only exploratory and further work is needed to establish better the link between these two groups of diseases. Further insights will possibly be gained from analyses in large clinical and administrative databases with the power to look at the relation between Alzheimer's disease and individual types of cancer. The potential impact of cancer treatment on risk of Alzheimer's disease is another interesting area for future work. Presently there are few curative treatments for cancer and not even one disease modifying drug for Alzheimer's disease. Our data suggest that vulnerability to cancer may actually protect against neurodegeneration, and vice versa. A further understanding of the basis for this inverse relation may lead to

novel therapies and should remain a focus of intense basic and translational research.

Contributors: JAD, AB, KPL, and DPK conceived and designed the study. BEK, GLS, and PAW acquired the data. JAD, AB, RA, and SS analysed and interpreted the data. JAD drafted the manuscript. AB, RA, BEK, GLS, TK, DPK, KPL, SS, and PAW critically revised the manuscript. AB carried out the statistical analysis. JAD and PAW obtained the funding. RA, BEK, GLS, TK, DPK, KPL, SS, and PAW supervised the study. JAD and AB had full access to all the data in the study. JAD takes responsibility for the integrity of the data and the accuracy of the data analysis and is guarantor.

Funding: JAD was funded by a Veterans' Administration career development award. The Framingham Heart Study is supported by an National Heart, Lung, and Blood Institute contract (N01-HC-25195) and by grants from the National Institutes of Aging (AG16495 and AG08122) and National Institute of Neurological Disorders and Stroke (NS17950). The authors' work was independent of the funders. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The Framingham study protocols and participant consent forms were approved by the institutional review board of Boston University School of Medicine.

Data sharing: No additional data available.

1 Corsellis J. Mental illness and the ageing brain. Oxford University Press, 1962.
2 DeSouky AL. The relationship between cancer and Alzheimer's disease. *J Am Geriatr Soc* 1992;40:1075.
3 Roe CM, Behrens MI, Xiong C, Miller JP, Morris JC. Alzheimer disease and cancer. *Neurology* 2005;64:895-8.
4 Roe CM, Fitzpatrick AL, Xiong C, Sieh W, Kuller L, Miller JP, et al. Cancer linked to Alzheimer disease but not vascular dementia. *Neurology* 2010;74:106-12.
5 Tirumalasetti F, Han L, Birkett DP. The relationship between cancer and Alzheimer's disease. *J Am Geriatr Soc* 1991;39:840.
6 Yamada M, Sasaki H, Mimori Y, Kasagi F, Sudoh S, Ikeda J, et al. Prevalence and risks of dementia in the Japanese population: RERF's adult health study of Hiroshima subjects. Radiation Effects Research Foundation. *J Am Geriatr Soc* 1999;47:189-95.
7 Bajaj A, Driver JA, Schernhammer ES. Parkinson's disease and cancer risk: a systematic review and meta-analysis. *Cancer Causes Control* 2010;21:697-707.
8 Driver JA, Kurth T, Buring JE, Gaziano JM, Logroscino G. Prospective case-control study of nonfatal cancer preceding the diagnosis of Parkinson's disease. *Cancer Causes Control* 2007;18:705-11.
9 Driver JA, Logroscino G, Buring JE, Gaziano JM, Kurth T. A prospective cohort study of cancer incidence following the diagnosis of Parkinson's disease. *Cancer Epidemiol Biomarkers Prev* 2007;16:1260-5.
10 Elbaz A, Peterson BJ, Bower JH, Yang P, Maraganore DM, McDonnell SK, et al. Risk of cancer after the diagnosis of Parkinson's disease: a historical cohort study. *Mov Disord* 2005;20:719-25.
11 Elbaz A, Peterson BJ, Yang P, Van Gerpen JA, Bower JH, Maraganore DM, et al. Nonfatal cancer preceding Parkinson's disease: a case-control study. *Epidemiology* 2002;13:157-64.
12 Olsen JH, Friis S, Frederiksen K, McLaughlin JK, Mellemkjaer L, Moller H. Atypical cancer pattern in patients with Parkinson's disease. *Br J Cancer* 2005;92:201-5.
13 Olsen JH, Tangerud K, Wermuth L, Frederiksen K, Friis S. Treatment with levodopa and risk for malignant melanoma. *Mov Disord* 2007;22:1252-7.
14 Behrens MI, Lendon C, Roe CM. A common biological mechanism in cancer and Alzheimer's disease? *Curr Alzheimer Res* 2009;6:196-204.
15 Driver JA, Lu KP. Pin1: a new genetic link between Alzheimer's disease, cancer and aging. *Curr Aging Sci* 2010;3:158-65.
16 Husseman JW, Nochlin D, Vincent I. Mitotic activation: a convergent mechanism for a cohort of neurodegenerative diseases. *Neurobiol Aging* 2000;21:815-28.
17 Lu KP. Pinning down cell signaling, cancer and Alzheimer's disease. *Trends Biochem Sci* 2004;29:200-9.
18 Lu KP, Liou YC, Vincent I. Proline-directed phosphorylation and isomerization in mitotic regulation and in Alzheimer's disease. *Bioessays* 2003;25:174-81.
19 Raina AK, Zhu X, Rottkamp CA, Monteiro M, Takeda A, Smith MA. Cyclin' toward dementia: cell cycle abnormalities and abortive oncogenesis in Alzheimer disease. *J Neurosci Res* 2000;61:128-33.
20 Staropoli JF. Tumorigenesis and neurodegeneration: two sides of the same coin? *Bioessays* 2008;30:719-27.
21 Wang W, Bu B, Xie M, Zhang M, Yu Z, Tao D. Neural cell cycle dysregulation and central nervous system diseases. *Prog Neurobiol* 2009;89:1-17.
22 West AB, Dawson VL, Dawson TM. To die or grow: Parkinson's disease and cancer. *Trends Neurosci* 2005;28:348-52.
23 Bennett DA. Is there a link between cancer and Alzheimer disease? *Neurology* 2010;75:1216-7.
24 Marwill SL, Freund KM, Barry PP. Patient factors associated with breast cancer screening among older women. *J Am Geriatr Soc* 1996;44:1210-4.
25 Scherder E, Bouma A, Borkent M, Rahman O. Alzheimer patients report less pain intensity and pain affect than non-demented elderly. *Psychiatry* 1999;62:265-72.
26 Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* 1951;41:279-81.
27 Kreger BE, Splansky GL, Schatzkin A. The cancer experience in the Framingham Heart Study cohort. *Cancer* 1991;67:1-6.
28 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
29 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939-44.
30 Centers for Disease Control and Prevention, Office on Smoking and Health. The health consequences of smoking: a report of the surgeon general. Centers for Disease Control and Prevention, 2004.
31 Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, et al. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. *N Engl J Med* 2002;346:476-83.
32 Seshadri S, Drachman DA, Lippa CF. Apolipoprotein E epsilon 4 allele and the lifetime risk of Alzheimer's disease. What physicians know, and what they should know. *Arch Neurol* 1995;52:1074-9.
33 Euser SM, Schram MT, Hofman A, Westendorp RG, Breteler MM. Measuring cognitive function with age: the influence of selection by health and survival. *Epidemiology* 2008;19:440-7.
34 Hernan MA, Takkouche B, Caamano-Isorna F, Gestal-Otero JJ. A meta-analysis of coffee drinking, cigarette smoking, and the risk of Parkinson's disease. *Ann Neurol* 2002;52:276-84.
35 Rusanen M, Kivipelto M, Quesenberry CP Jr, Zhou J, Whitmer RA. Heavy smoking in midlife and long-term risk of Alzheimer disease and vascular dementia. *Arch Intern Med* 2011;171:333-9.
36 Van Heemst D, Mooijaart SP, Beekman M, Schreuder J, de Craen AJ, Brandt BW, et al. Variation in the human TP53 gene affects old age survival and cancer mortality. *Exp Gerontol* 2005;40:11-5.
37 Lu KP, Hanes SD, Hunter T. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* 1996;380:544-7.
38 Lu KP, Zhou XZ. The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat Rev Mol Cell Biol* 2007;8:904-16.
39 Bao L, Sauter G, Sowadski J, Lu KP, Wang D. Prevalent overexpression of prolyl isomerase Pin1 in human cancers. *Am J Pathol* 2004;164:1727-37.
40 Liou Y-C, Sun A, Ryo A, Zhou XZ, Yu Z-X, Huang H-K, et al. Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration. *Nature* 2003;424:556-61.
41 Lu PJ, Wulf G, Zhou XZ, Davies P, Lu KP. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* 1999;399:784-8.
42 Sultana R, Boyd-Kimball D, Poon HF, Cai J, Pierce WM, Klein JB, et al. Oxidative modification and down-regulation of Pin1 in Alzheimer's disease hippocampus: a redox proteomics analysis. *Neurobiol Aging* 2006;27:918-25.
43 Wulf G, Garg P, Liou YC, Iglehart D, Lu KP. Modeling breast cancer in vivo and ex vivo reveals an essential role of Pin1 in tumorigenesis. *EMBO J* 2004;23:3397-407.
44 Lim J, Balastik M, Lee TH, Nakamura K, Liou YC, Sun A, et al. Pin1 has opposite effects on wild-type and P301L tau stability and tauopathy. *J Clin Invest* 2008;118:1877-89.
45 Segat L, Pontillo A, Annoni G, Trabattoni D, Vergani C, Clerici M, et al. PIN1 promoter polymorphisms are associated with Alzheimer's disease. *Neurobiol Aging* 2007;28:69-74.
46 Lu J, Hu Z, Wei S, Wang LE, Liu Z, El-Naggar AK, et al. A novel functional variant (-842G>C) in the PIN1 promoter contributes to decreased risk of squamous cell carcinoma of the head and neck by diminishing the promoter activity. *Carcinogenesis* 2009;30:1717-21.
47 Zotova E, Nicoll JA, Kalaria R, Holmes C, Boche D. Inflammation in Alzheimer's disease: relevance to pathogenesis and therapy. *Alzheimers Res Ther* 2010;2:1.
48 Webber KM, Raina AK, Marlatt MW, Zhu X, Prat MI, Morelli L, et al. The cell cycle in Alzheimer disease: a unique target for neuropharmacology. *Mech Ageing Dev* 2005;126:1019-25.
49 Seigers R, Fardell JE. Neurobiological basis of chemotherapy-induced cognitive impairment: a review of rodent research. *Neurosci Biobehav Rev* 2011;35:729-41.
50 Heck JE, Albert SM, Franco R, Gorin SS. Patterns of dementia diagnosis in surveillance, epidemiology, and end results breast cancer survivors who use chemotherapy. *J Am Geriatr Soc* 2008;56:1687-92.
51 Baxter NN, Durham SB, Phillips KA, Habermann EB, Virning BA. Risk of dementia in older breast cancer survivors: a population-based cohort study of the association with adjuvant chemotherapy. *J Am Geriatr Soc* 2009;57:403-11.
52 Surma-aho O, Niemela M, Vilkki J, Kouri M, Brander A, Salonen O, et al. Adverse long-term effects of brain radiotherapy in adult low-grade glioma patients. *Neurology* 2001;56:1285-90.

Accepted: 29 December 2011

Cite this as: *BMJ* 2012;344:e1442

What is already known on this topic

Parkinson's disease is associated with a decreased risk of most cancers, and growing evidence suggests that cancer and neurodegenerative diseases share genes and biological pathways

Limited data suggest an inverse relation between cancer and Alzheimer's disease

It is, however, unclear if the association might be result from selective mortality in cancer survivors or under-diagnosis in patients with Alzheimer's disease

What this study adds

Survivors of any cancer had a 33% lower risk of incident Alzheimer's disease and survivors of smoking related cancer had a 74% decreased risk, neither explained by survival bias

Participants with incident Alzheimer's disease had a 61% decreased risk of developing incident cancer, which may in part be due to under-diagnosis

This unusual pattern for cancer is similar to that seen in Parkinson's disease and suggests an inverse relation between cancer and neurodegeneration

Tables

Table 1| Characteristics of Framingham Heart Study participants who were free of dementia at examination 20 (baseline), by history of cancer. Values are numbers (percentages) unless stated otherwise

Characteristics	History of cancer (n=176)	No history of cancer (n=1102)
Mean (range) age (years)	77 (68-96)	76 (68-96)
Men	72 (41)	424 (38)
Women	104 (59)	678 (62)
Completion of secondary school	166 (72)	679 (67)
Apolipoprotein E4	38 (25)	200 (20)
Mean (range) homocysteine level (µmol/L)	12.9 (4.1-66.7)	12.8 (3.5-61.6)

Table 2| **Prevalent cancers at baseline and incident cancers diagnosed during follow-up**

		No (%) of cancers	
Cancer types	**Total No (n=423)**	**Prevalent (n=176)**	**Incident (n=247)**
Smoking related*	150	54 (31)	96 (39)
Non-smoking related	304	127 (72)	177 (72)
Cancers on routine screening	229	105 (60)	124 (50)
Head and neck	11	8 (5)	3 (1)
Oesophagus or stomach	23	9 (5)	14 (6)
Colon	49	21 (12)	28 (11)
Rectum	20	6 (3)	14 (6)
Pancreas	9	0 (0)	9 (4)
Lung	46	9 (5)	37 (15)
Haematological	15	2 (1)	13 (5)
Connective tissue	6	2 (1)	4 (2)
Melanoma	12	8 (5)	4 (2)
Breast	83	44 (25)	39 (16)
Uterus and endometrium	17	13 (7)	4 (2)
Cervix	6	5 (3)	1 (0)
Ovary	7	3 (2)	4 (2)
Prostate	59	21 (12)	38 (15)
Bladder	27	17 (10)	10 (4)
Kidney	10	4 (2)	6 (2)
Brain	5	0 (0)	5 (2)
Lymph nodes	7	3 (2)	4 (2)
Unknown primary	5	0 (0)	5 (2)
Other	6	1† (17)	5‡ (83)

*Some patients had both a smoking related and non-smoking related cancer, but only the first one was included in the analysis of overall cancer.
†Thyroid.
‡Liver, retroperitoneal, and pleural (n=1 each), gallbladder (n=2).

Table 3| Association between history of cancer at examination 20 (baseline) and incident dementia in Framingham Heart Study, after adjustment

	No of cancers		Hazard ratio (95% CI)*		
				Alzheimer's disease	
Model and cancer types	At baseline	Incident cases	Any dementia	Possible	Probable
Model 1 (n=1274)†:			n=322	n=256	n=220
All‡	175	247	0.83 (0.63 to 1.10)	0.81 (0.59 to 1.11)	0.67 (0.47 to 0.97)
Smoking related§	54	96	0.79 (0.45 to 1.39)	0.62 (0.31 to 1.26)	0.26 (0.08 to 0.82)
Non-smoking related	127	177	0.84 (0.62 to 1.13)	0.87 (0.62 to 1.21)	0.82 (0.57 to 1.19)
Model 2 (n=1037)¶:			n=263	n=212	n=185
All‡	133	210	0.92 (0.68 to 1.26)	0.90 (0.64 to 1.28)	0.76 (0.52 to 1.12)
Smoking related§	40	77	0.66 (0.32 to 1.33)	0.62 (0.28 to 1.41)	0.34 (0.11 to 1.08)
Non-smoking related	97	157	0.99 (0.72 to 1.38)	0.99 (0.68 to 1.42)	0.91 (0.61 to 1.35)

*Calculated using Cox proportional hazards modelling.

†Adjusted for age, sex, smoking, and incident cancer.

‡Does not include non-melanoma skin cancers.

§Defined as cancer of the oral cavity, pharynx, larynx, oesophagus, stomach, pancreas, lung, cervix, bladder, and kidney.

¶Additionally adjusted for apolipoprotein E4 status, education, and homocysteine level.

Table 4| Association between history of cancer at examination 20 (baseline) and incident dementia in Framingham Heart Study among 995 participants who survived to at least age 80

	No of cancers		Hazard ratio (95% CI)*		
			Any dementia (n=302)	Alzheimer's disease	
Cancer types	At baseline	Incident cases		Possible (n=244)	Probable (n=209)
All	127	183	0.87 (0.65 to 1.16)	0.81 (0.46 to 1.46)	0.88 (0.64 to 1.20)
Smoking related	37	69	0.83 (0.60 to 1.15)	0.68 (0.34 to 1.38)	0.88 (0.62 to 1.25)
Non-smoking related	94	128	0.68 (0.47 to 0.99)	0.29 (0.09 to 0.90)	0.83 (0.56 to 1.21)

*Calculated using Cox proportional hazards modelling, adjusted for age, sex, smoking, and incident cancer.

Table 5| **Distribution of incident cancer in participants with any dementia and matched controls**

Cancer type	Total No	No (%) of cases (n=495)	No (%) of controls (n=1485)
Overall	252	41 (8.3)	211 (14.2)
Non-smoking related	171	27 (5.5)	144 (9.7)
Smoking related	81	14 (2.8)	67 (4.5)
Cancers on routine screening	119	17 (3.4)	102 (6.9)
Colorectal	49	10 (2.0)	39 (2.6)
Lung	36	3 (0.6)	33 (2.2)
Breast	35	4 (0.8)	31 (2.1)
Prostate	27	2 (0.4)	25 (1.7)
Bladder	21	7 (1.4)	14 (0.9)
Haematological	12	2 (0.4)	10 (0.7)
Unknown primary	10	3 (0.6)	7 (0.5)
Melanoma	8	1 (0.2)	7 (0.5)
Pancreas	7	1 (0.2)	6 (0.4)
Kidney	6	0 (0.0)	6 (0.4)
Lymph nodes	5	1 (0.2)	4 (0.3)
Stomach	4	1 (0.2)	3 (0.2)
Corpus uteri	4	0 (0.0)	4 (0.3)
Head and neck	3	0 (0.0)	3 (0.2)
Liver	3	0 (0.0)	3 (0.2)
Male breast	3	1 (0.2)	2 (0.1)
Brain	3	1 (0.2)	2 (0.1)
Oesophagus	2	2 (0.4)	0 (0.0)
Gallbladder	2	0 (0.0)	2 (0.1)
Thyroid	2	0 (0.0)	2 (0.1)
Connective tissue	2	1 (50.0)	1 (50.0)
Cervix	2	0 (0.0)	2 (100.0)
Ovary	2	1 (50.0)	1 (50.0)
Small intestine	1	0 (0.0)	1 (100.0)
Pleura	1	0 (0.0)	1 (100.0)
Bone	1	0 (0.0)	1 (100.0)
Female genitalia	1	0 (0.0)	1 (100.0)

Table 6| Nested case-control study of confirmed dementia and incident cancer in Framingham Heart Study

	Any dementia (n=495)		Possible Alzheimer's disease (n=376)		Probable Alzheimer's disease (n=327)	
Baseline	**No/No***	**Hazard ratio (95% CI)†**	**No/No***	**Hazard ratio (95% CI)**	**No/No***	**Hazard ratio (95% CI)**
Any cancer:						
Model 1‡	252/1485	0.44 (0.32 to 0.61)	180/1128	0.38 (0.25 to 0.56)	159/981	0.39 (0.26 to 0.58)
Model 2§	175/913	0.38 (0.26 to 0.57)	123/696	0.29 (0.17 to 0.49)	111/599	0.29 (0.17 to 0.49)
Smoking related cancer:						
Model 1	88/1485	0.45 (0.26 to 0.77)	66/1128	0.45 (0.24 to 0.84)	58/981	0.45 (0.24 to 0.88)
Model 2	61/913	0.31 (0.15 to 0.65)	46/696	0.24 (0.10 to 0.61)	40/599	0.21 (0.07 to 0.58)
Non-smoking related cancer:						
Model 1	178/1485	0.45 (0.31 to 0.65)	126/1128	0.36 (0.22 to 0.58)	110/981	0.36 (0.21 to 0.59)
Model 2	125/913	0.42 (0.26 to 0.67)	86/696	0.32 (0.17 to 0.59)	78/599	0.31 (0.16 to 0.59)

*Number of cancer cases/number of reference participants.

†Hazard ratio and 95% confidence interval calculated using Cox proportional hazards modelling.

‡Matched on age and sex.

§Additionally adjusted for body mass index and smoking.

Binge Drinking Intensity
A Comparison of Two Measures

Marissa B. Esser, MPH, Dafna Kanny, PhD, Robert D. Brewer, MD, MSPH, Timothy S. Naimi, MD, MPH

Background: Binge drinking (≥4 drinks for women; ≥5 drinks for men, per occasion) is responsible for more than half of the estimated 80,000 U.S. deaths annually and three-quarters of the $223.5 billion in costs in 2006. Binge drinking prevalence is assessed more commonly than binge drinking intensity (i.e., number of drinks consumed per binge episode). Risk of binge drinking–related harm increases with intensity, and thus it is important to monitor. The largest number of drinks consumed is assessed in health surveys, but its usefulness for assessing binge intensity is unknown.

Purpose: To assess the agreement between two potential measures of binge drinking intensity: the largest number of drinks consumed by binge drinkers (maximum-drinks) and the total number of drinks consumed during their most recent binge episode (drinks-per-binge).

Methods: Data were analyzed from 7909 adult binge drinkers from 14 states responding to the 2008 Behavioral Risk Factor Surveillance System (BRFSS) binge drinking module. Mean and median drinks-per-binge from that module were compared to mean and median maximum-drinks. Analyses were conducted in 2010–2011.

Results: Mean (8.2) and median (5.9) maximum-drinks were strongly correlated with mean (7.4) and median (5.4) drinks-per-binge ($r=0.57$). These measures were also strongly correlated across most sociodemographic and drinking categories overall and within states.

Conclusions: The maximum-drinks consumed by binge drinkers is a practical method for assessing binge drinking intensity and thus can be used to plan and evaluate Community Guide–recommended strategies for preventing binge drinking (e.g., increasing the price of alcoholic beverages and regulating alcohol outlet density).

(Am J Prev Med 2012;42(6):625–629) Published by Elsevier Inc. on behalf of American Journal of Preventive Medicine

Introduction

Excessive alcohol consumption is responsible for an average of 80,000 deaths in the U.S. each year[1] and $223.5 billion in economic costs in 2006.[2] More than half of these deaths and three-quarters of the economic costs are due to binge drinking[1,2] (≥4 drinks for women; ≥5 drinks for men, per occasion).[3,4] Binge drinking also is associated with a range of health and social problems, such as motor vehicle crashes, interpersonal violence, new HIV infections and sexually transmitted infections, liver cirrhosis, cancers, stroke, and alcohol dependence.[5–8]

The risk of binge drinking–related harm increases with the intensity of binge drinking (i.e., the number of drinks consumed).[9] Yet, adult binge drinkers often report drinking at levels that far exceed those used to define this pattern of alcohol consumption.[10,11] It is therefore important to routinely monitor binge drinking intensity to assess the public health impact of this behavior and to plan and evaluate evidence-based strategies to prevent it. However, questions on the number of drinks consumed by binge drinkers are not asked routinely in state health risk behavior surveys, such as the Behavioral Risk Factor Surveillance System (BRFSS). More commonly, health surveys include questions on the largest number of drinks consumed in a given time period (i.e., maximum-drinks),[12,13] but it is not clear whether these questions are useful for assessing binge drinking intensity.

From the Alcohol Program, Epidemiology and Surveillance Branch (Esser, Kanny, Brewer), Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion, CDC, Atlanta, Georgia; and the Section of General Internal Medicine (Naimi), Boston University Medical Center, Boston, Massachusetts

Address correspondence to: Dafna Kanny, PhD, Alcohol Program, National Center for Chronic Disease Prevention and Health Promotion, 4770 Buford Hwy NE, MS-K67, Atlanta GA 30341. E-mail: dkanny@cdc.gov.

0749-3797/$36.00

doi: 10.1016/j.amepre.2012.03.001

The purpose of the present study was to assess the agreement between the following two measures of binge drinking intensity: the mean and median maximum-drinks reported by binge drinkers on any occasion and mean and median total drinks consumed by binge drinkers during their most recent binge episode (i.e., drinks-per-binge). If there is strong agreement between these two measures, maximum-drinks may be a way to routinely measure state-specific binge drinking intensity, thus improving public health surveillance on the number of drinks consumed by adult binge drinkers.

Methods

Data came from the 2008 BRFSS. Sampling design, purpose, and analysis descriptions are available at www.cdc.gov/brfss/. Binge drinkers were defined as current drinkers who reported consuming ≥4 drinks (women) or ≥5 drinks (men) on an occasion during the past 30 days in the BRFSS core survey. Maximum-drinks consumed was assessed in the core survey by the largest number of drinks consumed on any occasion during the past 30 days. In 2008, a total of 14 states administered an optional module to obtain information about binge drinkers' most recent binge episode, including the number of drinks consumed by beverage type. Drinks-per-binge was calculated by summing beverage-specific consumption (i.e., beer, wine, liquor, and pre-mixed flavored drinks) from that episode.

Analyses were restricted to binge drinkers from the 14 states who administered the 2008 binge drinking module. Respondents with missing information on beverage-specific consumption, or those who reported maximum-drinks or drinks-per-binge that were <4 for women or <5 for men were excluded. The final sample consisted of 7909 respondents.

Data analysis was conducted in 2010–2011 using SAS-callable SUDAAN 9.2. Results were weighted by age, gender, and race/ethnicity to be representative of people residing in the states and to the respondent's selection probability. Means and 95% CIs, medians, and Pearson correlation coefficients of the maximum-drinks and drinks-per-binge were assessed by sociodemographic characteristics and by state. Significant differences between means were determined by non-overlapping CIs.

Results

The study population of binge drinkers was predominantly male (58.0%); aged 18–54 years (76.9%); non-Hispanic white (83.3%); had at least some college education (63.9%); and had household incomes of ≥$50,000 (52.6%). There was a strong correlation between maximum-drinks (M=8.2; median=5.9) and drinks-per-binge (M=7.4; median=5.4) (r=0.57, p<0.001) (Table 1). These measures were strongly correlated for binge drinkers across most groups, except those with less than high school education (r=0.29, p<0.001); Hispanic race/ethnicity (r=0.38, p<0.001); and those who reported ≥5 binge drinking episodes in the past month (r=0.45, p<0.001). Mean and median maximum-drinks and drinks-per-binge were also strongly correlated in all states except Georgia (Table 2).

Discussion

Overall, binge drinkers' maximum-drinks correlated strongly with drinks-per-binge, an established measure of binge drinking intensity.[10] Measures generally correlated strongly within individual states as well. Further, the mean and median drinks-per-binge were approximately 90% of the mean and median maximum-drinks, respectively, further supporting the level of agreement between these two measures of binge drinking intensity. To our knowledge, this is the first study to assess the agreement between the largest number of drinks consumed by binge drinkers (maximum-drinks) and the number of drinks consumed during the most-recent binge drinking episode (drinks-per-binge) as measures of binge drinking intensity.

The findings of the current study have important implications for public health surveillance on binge drinking among adults because, unlike the drinks-per-binge measure, the maximum-drinks measure is assessed in all states annually. Therefore, it can be used routinely to assess binge intensity, and to plan and evaluate evidence-based binge drinking prevention strategies in states. Although the correlation between maximum-drinks and drinks-per-binge was generally quite strong across sociodemographic groups and within states, the correlations between these measures were weaker for respondents of Hispanic race/ethnicity and with less than high school degree, even though the mean and median maximum-drinks and drinks-per-binge within these strata were quite similar. This inconsistency may be due to the smaller sample sizes in these strata.

The greater binge drinking intensity of those reporting ≥5 binge drinking episodes/month relative to those reporting 1–2 episodes is consistent with other studies.[10] Among respondents reporting ≥5 episodes/month, the higher mean and median maximum-drinks than drinks-per-binge may reflect greater variability in binge intensity among frequent binge drinkers. Frequent binge drinkers (≥5 episodes/month) had more episodes in which the maximum-drinks could exceed the drinks-per-binge measure. Consequently, there was a greater likelihood that the most recent binge episode was not the most intense. However, correlations between binge intensity measures were still strong for those who reported ≥5 episodes/month, supporting the use of maximum-drinks to assess binge drinking intensity among frequent binge drinkers too.

The present study has limitations. First, restricting the sample to those who reported alcohol consumption at or above the binge drinking threshold on multiple measures may have reduced the generalizability of the findings. However, the binge intensity estimates were not substan-

Table 1. Maximum-drinks among adult binge drinkers and drinks-per-binge by sociodemographic characteristics and binge drinking episode frequency[a]

Characteristics	*n*	Maximum-drinks[b]		Drinks-per-binge[c]		Pearson correlation coefficient[d]
		M (95% CI)	Median	M (95% CI)	Median	
Overall	7909	8.2 (7.9, 8.4)	5.9	7.4 (7.1, 7.6)	5.4	0.57
Gender						
Male	4588	9.2 (8.9, 9.6)	7.1	8.2 (7.9, 8.6)	5.9	0.53
Female	3321	5.9 (5.7, 6.0)	4.7	5.6 (5.5, 5.8)	4.2	0.53
Age group (years)						
18–34	2107	9.1 (8.6, 9.6)	6.9	8.1 (7.6, 8.6)	5.8	0.52
35–54	3971	7.5 (7.2, 7.8)	5.6	6.9 (6.7, 7.2)	5.3	0.64
≥55	1816	6.6 (6.3, 6.9)	5.1	6.1 (5.9, 6.4)	4.8	0.52
Race/ethnicity						
Non-Hispanic white	6589	7.7 (7.5, 8.0)	5.8	7.1 (6.7, 7.2)	5.2	0.69
Hispanic	653	9.2 (8.5, 9.9)	7.2	8.6 (7.8, 9.5)	6.0	0.38
Other[e]	630	9.0 (7.6, 10.5)	5.8	7.8 (7.1, 8.5)	5.7	0.50
Education						
<high school graduate	488	9.9 (8.6, 11.2)	7.5	9.0 (7.8, 10.3)	6.6	0.29
High school graduate	2368	8.9 (8.3, 9.4)	6.7	8.1 (7.5, 8.7)	5.9	0.62
Some college	2365	8.1 (7.7, 8.6)	5.9	7.1 (6.8, 7.4)	5.3	0.62
College graduate	2686	7.1 (6.7, 7.5)	5.5	6.6 (6.3, 7.0)	5.0	0.64
Income ($)						
0–<49,999	3372	8.9 (8.4, 9.3)	6.6	8.0 (7.5, 8.5)	5.7	0.56
≥50,000	4162	7.7 (7.3, 8.0)	5.7	7.0 (6.7, 7.3)	5.2	0.58
Binge episodes						
≥5	1869	11.2 (10.5, 12.0)	9.3	8.9 (8.2, 9.6)	6.4	0.45
3–4	1451	8.4 (7.9, 8.8)	6.7	7.8 (7.3, 8.2)	5.9	0.62
1–2	4589	6.5 (6.3, 6.7)	5.2	6.5 (6.2, 6.7)	5.0	0.65

[a]Data were included from the 14 states that conducted the optional binge drinking module: Alaska, California, Delaware, Georgia, Iowa, Maine, Michigan, Montana, Nebraska, Nevada, New Mexico, Texas, Wisconsin, and Wyoming. All analyses were restricted to include only binge drinkers.

[b]Largest number of drinks consumed on any occasion during the past 30 days, among binge drinkers; data are from the BRFSS core section.

[c]Total number of drinks consumed in most recent binge drinking episode in the past 30 days; binge drinking was defined as consuming ≥4 drinks for women and ≥5 drinks for men per occasion in the past 30 days. Data are from the BRFSS binge drinking module.

[d]Correlation between maximum-drinks among binge drinkers and drinks-per-binge, within respective categories of sociodemographics and drinking patterns; all *p*-values were significant at <0.001.

[e]Non-Hispanic black, Asian, Native Hawaiian/Pacific Islander, American Indian/Alaskan Native, multiracial, and other

tially different from those reported in other publications and it therefore seems unlikely that the restricted study population substantially changed the relationship between the two evaluated binge drinking intensity measures.[10,11] Second, self-reported estimates of binge drinking intensity are likely to be underestimated because of recall bias[14]; social desirability response bias; nonresponse bias[15]; and because of the increasing number of cell phone–only households, particularly among young adults.[16,17] In fact, a recent study found that BRFSS alcohol consumption data accounted for a median of 22%–32% of state consumption based on alcohol sales.[18] Therefore, it seems unlikely that either of the measures of binge drinking intensity that were assessed in the current study overestimated the actual number of drinks consumed by adult binge drinkers in the U.S.

Table 2. Maximum-drinks among binge drinkers and drinks-per-binge by state

State	*n*	Maximum-drinks[a]		Drinks-per-binge[b]		Pearson correlation coefficient[c]
		M (95% CI)	Median	M (95% CI)	Median	
Overall	7909	8.2 (7.9, 8.4)	5.9	7.4 (7.1, 7.6)	5.4	0.56
Alaska	313	7.3 (6.8, 7.8)	5.7	7.0 (6.6, 7.4)	5.5	0.70
California	516	8.1 (7.3, 8.7)	6.0	7.5 (6.9, 8.1)	5.3	0.58
Delaware	378	7.4 (6.9, 7.8)	5.7	7.3 (6.8, 7.8)	5.4	0.73
Georgia	393	8.1 (6.8, 9.3)	5.6	7.9 (7.1, 8.7)	5.8	0.33
Iowa	741	7.7 (7.4, 8.1)	6.1	6.9 (6.5, 7.2)	5.3	0.60
Maine	385	7.9 (7.3, 8.6)	5.7	7.5 (6.8, 8.3)	5.6	0.74
Michigan	299	7.9 (7.3, 8.5)	5.9	7.0 (6.6, 7.4)	5.5	0.56
Montana	726	8.1 (7.6, 8.6)	5.9	6.8 (6.3, 7.3)	4.8	0.60
Nebraska	596	7.3 (6.8, 7.8)	5.6	6.7 (6.3, 7.1)	5.3	0.48
Nevada	567	8.6 (7.6, 9.6)	5.7	8.3 (7.3, 9.3)	5.5	0.62
New Mexico	381	7.7 (7.0, 8.3)	5.7	7.2 (6.6, 7.8)	5.4	0.61
Texas	798	8.6 (8.0, 9.2)	6.0	7.3 (6.8, 7.7)	5.4	0.65
Wisconsin	1054	8.3 (7.9, 8.7)	6.4	7.4 (7.0, 7.7)	5.6	0.65
Wyoming	762	8.4 (7.8, 8.9)	5.9	7.9 (7.5, 8.4)	5.9	0.71

[a]Largest number of drinks consumed on any occasion during the past 30 days, among binge drinkers; data are from the BRFSS core section.
[b]Total number of drinks consumed in most recent binge drinking episode in the past 30 days; binge drinking was defined as consuming ≥4 drinks for women and ≥5 drinks for men per occasion in the past 30 days. Data are from the BRFSS binge drinking module.
[c]Correlation between maximum-drinks among binge drinkers and drinks-per-binge, within states; all *p*-values were significant at <0.001.

The present study affirms that binge drinkers generally drink at levels that are well above those used to define this behavior. Therefore, in addition to monitoring binge drinking prevalence and frequency, it is important to monitor binge intensity in states and nationwide, which routinely can be done with the maximum-drinks measure. This information can, in turn, be used to plan and evaluate strategies for preventing binge drinking and related harms, such as those recommended by the *Guide to Community Preventive Services* (e.g., increasing the price of alcoholic beverages and regulating alcohol outlet density).[19–21]

The authors thank the BRFSS coordinators from the 14 states that conducted the binge drinking module and members of the Survey Operation Team in the Division of Behavioral Surveillance; Public Health Surveillance Program Office; Office of Surveillance, Epidemiology and Laboratory Services; and CDC for their help in collecting the data used in this study.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the CDC.

No financial disclosures were reported by the authors of this paper.

References

1. CDC. Alcohol-related disease impact (ARDI) application. apps.nccd.cdc.gov/DACH_ARDI/Default.aspx.
2. Bouchery EE, Harwood HJ, Sacks JJ, Simon CJ, Brewer RD. Economic costs of excessive alcohol consumption in the U.S., 2006. Am J Prev Med 2011;41(5):516–24.
3. National Institute on Alcohol Abuse and Alcoholism. NIAAA Council approves binge drinking definition. NIAAA Newslett 2004;(3):3.
4. Wechsler H, Dowdall GW, Davenport A, Rimm EB. A gender-specific measure of binge drinking among college students. Am J Public Health 1995;85(7):982–5.
5. Hingson RW, Zha W. Age of drinking onset, alcohol use disorders, frequent heavy drinking, and unintentionally injuring oneself and others after drinking. Pediatrics 2009;123(6):1477–84.
6. Hingson RW, Edwards EM, Heeren T, Rosenbloom D. Age of drinking onset and injuries, motor vehicle crashes, and physical fights after drinking and when not drinking. Alcohol Clin Exp Res 2009;33(5):783–90.
7. Naimi TS, Brewer RD, Mokdad A, Denny C, Serdula MK, Marks JS. Binge drinking among U.S. adults. JAMA 2003;289(1):70–5.

8. Chikritzhs TN, Jonas HA, Stockwell TR, Heale PF, Dietze PM. Mortality and life-years lost due to alcohol: a comparison of acute and chronic causes. Med J Aust 2001;174(6):281–4.
9. Wechsler H, Nelson TF. Relationship between level of consumption and harms in assessing drink cut-points for alcohol research: commentary on "Many college freshmen drink at levels far beyond the binge threshold" by White et al. Alcohol Clin Exp Res 2006;30(6):922–7.
10. Naimi TS, Nelson DE, Brewer RD. The intensity of binge alcohol consumption among U.S. adults. Am J Prev Med 2010;38(2):201–7.
11. CDC. Vital signs: binge drinking prevalence, frequency, and intensity among adults—U.S., 2010. MMWR Morb Mortal Wkly Rep 2012; 61(1):14–9.
12. Greenfield TK, Nayak MB, Bond J, Ye Y, Midanik LT. Maximum quantity consumed and alcohol-related problems: assessing the most alcohol drunk with two measures. Alcohol Clin Exp Res 2006;30(9):1576–82.
13. CDC. Behavioral Risk Factor Surveillance System (BRFSS) survey questionnaire, 2005–present. Atlanta GA. www.cdc.gov/brfss/questionnaires/english.htm.
14. Gmel G, Daeppen JB. Recall bias for seven-day recall measurement of alcohol consumption among emergency department patients: implications for case-crossover designs. J Stud Alcohol Drugs 2007;68(2):303–10.
15. Stockwell T, Donath S, Cooper-Stanbury M, Chikritzhs T, Catalano P, Mateo C. Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall. Addiction 2004;99(8):1024–33.
16. Delnevo CD, Gundersen DA, Hagman BT. Declining estimated prevalence of alcohol drinking and smoking among young adults nationally: artifacts of sample undercoverage? Am J Epidemiol 2008; 167(1):15–9.
17. Blumberg SJ, Luke JV, Ganesh N, Davern ME, Boudreaux MH, Soderberg K. Wireless substitution: State-level estimates from the National Health Interview Survey, January 2007–June 2010. Natl Health Stat Rep 2011;(39):1–26.
18. Nelson DE, Naimi TS, Brewer RD, Roeber J. U.S. state alcohol sales compared to survey data, 1993-2006. Addiction 2010;105(9): 1589–96.
19. Task Force on Community Prevention Services. Preventing excessive alcohol consumption. www.thecommunityguide.org/alcohol.
20. Elder RW, Lawrence B, Ferguson A, et al. Task Force on Community Preventive Services. The effectiveness of tax policy interventions for reducing excessive alcohol consumption and related harms. Am J Prev Med 2010;38(2):217–29.
21. Task Force on Community Preventive Services. Increasing alcohol beverage taxes is recommended to reduce excessive alcohol consumption and related harms. Am J Prev Med 2010;38(2):230–2.

American Journal of Epidemiology

Vol. 175, No. 4
DOI: 10.1093/aje/kwr318
Advance Access publication:
January 6, 2012

Original Contribution

Mortality Among Young Injection Drug Users in San Francisco: A 10-Year Follow-up of the UFO Study

Jennifer L. Evans*, Judith I. Tsui, Judith A. Hahn, Peter J. Davidson, Paula J. Lum, and Kimberly Page

* Correspondence to Jennifer L. Evans, Prevention and Public Health Group, Department of Epidemiology and Biostatistics, University of California, San Francisco, 50 Beale Street, Suite 1200, San Francisco, CA 94105 (e-mail: jevans@epi.ucsf.edu).

Initially submitted March 18, 2011; accepted for publication August 16, 2011.

This study examined associations between mortality and demographic and risk characteristics among young injection drug users in San Francisco, California, and compared the mortality rate with that of the population. A total of 644 young (<30 years) injection drug users completed a baseline interview and were enrolled in a prospective cohort study, known as the UFO ("U Find Out") Study, from November 1997 to December 2007. Using the National Death Index, the authors identified 38 deaths over 4,167 person-years of follow-up, yielding a mortality rate of 9.1 (95% confidence interval: 6.6, 12.5) per 1,000 person-years. This mortality rate was 10 times that of the general population. The leading causes of death were overdose (57.9%), self-inflicted injury (13.2%), trauma/accidents (10.5%), and injection drug user-related medical conditions (13.1%). Mortality incidence was significantly higher among those who reported injecting heroin most days in the past month (adjusted hazard ratio = 5.8, 95% confidence interval: 1.4, 24.3). The leading cause of death in this group was overdose, and primary use of heroin was the only significant risk factor for death observed in the study. These findings highlight the continued need for public health interventions that address the risk of overdose in this population in order to reduce premature deaths.

drug users; epidemiology; hepatitis C; mortality; overdose; young adult

Abbreviations: CI, confidence interval; HBV, hepatitis B virus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; ICD, *International Classification of Diseases*; IDU, injection drug user; IQR, interquartile range; IRR, incidence rate ratio; NDI, National Death Index; UFO, "U Find Out."

Injection drug use has been associated with excess morbidity and mortality (1, 2). A number of studies have reported incidence that greatly exceeds that of the general population (3–5). Risk factors that have been associated with mortality among injection drug users (IDUs) from prior studies include infectious diseases, such as human immunodeficiency virus (HIV) and hepatitis C virus (HCV) (3, 6–8). Given improvements in medical care for HIV and HCV as well as more widespread implementation of public health interventions for IDUs (needle exchange, overdose prevention, opiate substitution treatment, and so on), current studies are needed to assess whether the incidence and causes of mortality have shifted over time. Additionally, studies of unique subpopulations such as young IDUs are needed to understand the patterns and risk factors for mortality in order to inform future tailored interventions.

The current study was undertaken to assess the incidence of mortality and risk factors for mortality among young injection drug users in San Francisco, California. Given that the UFO ("U Find Out") Study has been collecting detailed information on young IDUs since the late 1990s, it provides a unique opportunity to study mortality trends and risk factors for mortality in this population. The specific aims of the study were as follows: 1) to investigate trends in mortality incidence overall and overdose-specific mortality, 2) to calculate standardized mortality ratios, and 3) to identify risk factors for mortality in young injectors in San Francisco between 1997 and 2007.

MATERIALS AND METHODS

Beginning in 1997, young (<30 years) IDUs in San Francisco have been offered participation in multiple prospective studies under variations of the shared title known as the UFO Study and described previously (9–11). In brief, young IDUs were recruited by peer outreach workers familiar with neighborhoods in San Francisco where young IDUs congregate, using study invitation cards and flyers, contacts with youth friendly neighborhood groups and community providers, and word of mouth. Inclusion criteria for screening were the following: 1) less than 30 years of age; 2) self reported use of injection drugs in the past 30 days, 3) ability to provide informed consent, 4) understanding spoken English and, after 2005, 5) self-reported HCV-negative or unknown status. The UFO Study had 3 waves of data collection between November 1997 and December 2007. In the first wave from 1997 to 1999, subjects completed a baseline screening interview and underwent counseling and serologic testing for HIV, hepatitis B, and hepatitis C virus. Subjects without evidence of acute or chronic hepatitis B virus (HBV) infection or immunization to HBV were eligible for enrollment in a prospective study of HBV vaccination (UFO-2 Study). Starting in 2000, subjects who were HCV negative at the baseline screening were eligible for enrollment into the UFO-3 Study. The UFO-3 Study had 2 waves of recruitment, from 2000 to 2002 and from 2003 to 2007. Participants were included in this analysis only if they were enrolled prior to December 31, 2007, the end date for searching for death records. All research protocols and informed consent were approved by the Institutional Review Board of the University of California, San Francisco.

Survey instrument and measures

Eligible consenting participants were interviewed, counseled, and tested for antibodies to HCV (anti-HCV) and presence of viremia (HCV RNA) at baseline. Follow-up included monthly ''check-ins'' and quarterly study visits that included structured interviews to assess risk (principally drug-related and sexual) exposures, HCV status (including anti-HCV and HCV RNA testing), and risk reduction counseling and referrals. In the first and second study waves, testing for antibodies to HIV and HBV was performed at baseline and follow-up visits. Deaths were ascertained though the National Death Index (NDI) through the end of 2007. Methods for establishing a match are described in detail on the NDI Web site (http://www.csc.gov/nchs/ndi.htm). In brief, records are matched by first and last name, date of birth, sex, and state of birth. Causes of death were obtained from the NDI-Plus. The underlying cause of death was based on the *International Classification of Diseases* (ICD), Tenth Revision. The primary outcome for this analysis was date of death as identified by an NDI match. Causes of death were reviewed, and overdose death was defined as ICD, Tenth Revision, codes X42, X44, X62, and Y14. Sociodemographic variables included in the analyses were age, gender, race/ethnicity (Caucasian vs. non-Caucasian), educational level (less than high school vs. high school or greater), recent (past 3 months) homelessness, and incarceration. Starting in wave 2, participants were asked about participation in drug treatment programs in the past 3 months and in the past week in wave 3. Recent drug treatment programs included drug detoxification, residential treatment, methadone or buprenorphine maintenance, and 12-step programs. Risk-related exposures included duration of injection drug use; frequency of injection; injection of heroin, cocaine, methamphetamine, crack, or heroin mixed with methamphetamine or cocaine; noninjection crack use; and nonfatal heroin overdose. Overdose was defined as a loss of consciousness where at least one intervention was attempted by a third party.

Data analysis

We calculated mortality incidence overall and by demographic and risk characteristics reported at the baseline interview. Standardized mortality ratios were calculated in 3-year intervals adjusting for age, sex, and race by using national mortality statistics from the National Center for Health Statistics. Standardized mortality ratios were calculated overall, separately for males and females, and for overdose-related deaths only; 95% confidence intervals were calculated by assuming that the observed deaths followed a Poisson distribution. Cox regression models with time-varying covariates were used to identify predictors of mortality. Survival time was defined as the time from initiation of injection (self-reported as age at first injected drugs) to death. Given that subjects were current injectors at the start of the study, data were left truncated. Subjects entered into the analysis at the baseline visit, and they remained until the date of death or were censored at December 31, 2007, or the last interview date if the last interview was after December 31, 2007. Covariates were included in the multivariate model if the incidence rate ratio for mortality reached statistical significance at a level of $P < 0.10$ in the bivariate analysis or were potential confounders (study wave, age, gender, HCV status, and duration of injection drug use). Baseline characteristics including age, gender, duration of injection, and study wave were entered into the model as fixed covariates. Drug use and injection-risk behaviors, overdose, and HCV status were entered as time-varying covariates. All analyses were conducted with the Stata statistical software package (release 11.2; StataCorp LP, College Station, Texas).

RESULTS

A total of 644 participants completed a baseline interview and were enrolled in the prospective cohort from November 1997 to December 2007. The median age was 22.0 years (interquartile range (IQR): 19.8–25.0), and participants had injected for a median of 4.0 years (IQR: 1.7–6.7) at the baseline interview. Sixty-eight percent of participants were male, and 78% were Caucasian (Table 1); 47% had less than a high school education, and 68% reported being homeless or marginally housed. At the most recent interview, the drug most often injected in the past 30 days was heroin (63%) followed by methamphetamine (27%); 31% reported injecting every day in the past month, 68% used a syringe exchange in the past month, and 32% were anti-HCV positive.

We identified 38 deaths over the follow-up period through 2007 using the NDI. The overall mortality rate over 4,167

Table 1. Selected Baseline Sociodemographic and Behavioral Characteristics at Last Quarterly Follow-up Visit for Young Injection Drug Users in the UFO Study, San Francisco, California, 1997–2007

Characteristic	No.	%	Proportion Mortality, %	No. of Person-Years	Mortality Rate/1,000 Person-Years	Incidence Rate Ratio	95% CI
Overall	644	100.0	5.9	4,167.36	9.12		
Study wave							
1997–1999	133	20.7	11.3	1,148.92	13.06	1.00	Referent
2000–2002	334	51.9	5.4	2,395.66	7.51	0.55	0.27, 1.14
2003–2007	177	27.5	2.8	622.78	8.03	0.80	0.51, 1.27
Age, years							
15–19	164	25.5	6.7	1,113.37	9.88	1.00	Referent
20–24	292	45.3	5.1	1,902.58	7.88	0.79	0.36, 1.76
25–30	188	29.2	6.4	1,151.41	10.42	1.03	0.68, 1.55
Gender							
Female	204	31.7	4.4	1,383.47	6.51	1.00	Referent
Male	440	38.3	6.6	2,783.89	10.42	1.60	0.76, 3.38
Education							
Less than high school	300	47.0	4.3	2,009.27	6.47	1.00	Referent
High school or more	339	53.0	7.1	2,136.13	11.24	1.74	0.88, 3.41
Race/ethnicity							
Caucasian	502	78.1	5.6	3,261.45	8.45	1.00	Referent
Non-Caucasian	141	21.9	7.1	842.90	11.78	1.39	0.71, 3.17
Homeless, past 3 months							
No	207	32.2	4.4	1,322.38	6.81	1.00	Referent
Yes	436	67.8	6.7	2,840.29	10.21	1.50	0.71, 3.17
Sexual behavior							
Female	204	31.7	4.5	1,383.47	6.51	1.00	Referent
Heterosexual male	261	40.5	8.4	1,710.26	12.86	1.89	0.93, 3.83
Men who have sex with men	179	27.8	3.9	1,074.62	6.52	1.00	0.61, 1.64
Hepatitis C virus positive							
No	433	67.9	4.9	2,782.30	7.55	1.00	Referent
Yes	205	32.1	8.3	1,356.27	12.53	1.66	0.88, 3.15
Human immunodeficiency virus positive							
No	565	95.8	6.4	3,780.08	9.52		
Yes	25	4.2	0.0	149.89	0.00		
Age of first drug injection, years							
<17	243	37.7	6.6	1,589.72	10.06	1.00	Referent
17–19	217	33.7	6.0	1,452.42	8.95	0.89	0.43, 1.84
≥20	184	28.6	4.9	1,125.22	8.00	0.89	0.60, 1.33

Table continues

person-years of follow-up was 9.12 (95% confidence interval (CI): 6.63, 12.53) per 1,000 person-years of observation. In females, the mortality rate was 6.51/1,000 person-years of observation and, in males, 10.42/1,000 person-years of observation. The median age at death was 26 years (IQR: 23–30), and the median time from initiation of injecting until death was 7.7 years (IQR: 4.9–11.3).

Increased mortality was observed among those who identified heroin or heroin mixed with other drugs as the drug injected most days in the 30 days prior to the most recent interview, relative to those who injected methamphetamine (incidence rate ratio (IRR) = 3.03, 95% CI: 1.47, 6.23) (Table 1). The incidence of mortality was significantly higher in those who reported having an overdose in the past 3 months compared with those who did not (IRR = 2.71, 95% CI: 1.35, 5.47). Mortality rates were elevated among HCV-seropositive individuals compared with negatives but did not reach statistical significance (IRR = 1.66,

Table 1. Continued

Characteristic	No.	%	Proportion Mortality, %	No. of Person-Years	Mortality Rate/1,000 Person-Years	Incidence Rate Ratio	95% CI
Duration injecting, years							
≤2	194	30.2	5.2	1,281.74	7.80	1.00	Referent
>2–5	209	32.5	5.3	1,377.33	7.99	1.02	0.44, 2.41
>5	240	37.3	7.1	1,503.37	11.31	1.20	0.82, 1.75
Injected every day, past month							
No	446	69.3	5.2	2,840.64	8.10	1.00	Referent
Yes	198	30.7	7.6	1,326.72	11.31	1.40	0.73, 2.68
Drug injected most days, past month							
Speed/methamphetamine	168	26.5	1.2	1,100.15	1.82	1.00	Referent
Heroin/heroin mix	441	69.6	7.9	2,900.47	12.07	3.03*	1.47, 6.23
Other	25	3.9	4.0	137.45	7.28	3.08	0.51, 18.66
Injected alone, past 3 months[a]							
No	338	53.7	5.0	2,184.05	7.78	1.00	Referent
Yes	292	46.3	7.2	1,879.04	11.18	1.44	0.76, 2.72
Syringe exchange, past month							
No	203	32.0	5.4	1,270.33	8.66	1.00	Referent
Yes	432	68.0	6.3	2,833.50	9.53	1.10	0.55, 2.22
Ever overdosed							
No	391	60.9	4.9	2,519.38	7.54	1.00	Referent
Yes	251	39.1	7.6	1,638.79	11.59	1.54	0.81, 2.90
Overdose, past 3 months[a]							
No	563	88.1	4.8	3,595.02	7.51	1.00	Referent
Yes	76	11.9	14.5	539.66	20.38	2.71*	1.35, 5.47
Incarcerated, past 3 months[a]							
No	404	63.0	559	2,561.85	8.59	1.00	Referent
Yes	237	37.0	6.8	1,586.35	10.09	1.17	0.62, 2.24
Drug treatment, past 3 months ($n = 504$)[b]							
No	379	75.2	4.0	2,218.70	6.76	1.00	Referent
Yes	125	24.8	6.4	754.29	10.61	1.57	0.67, 3.70

Abbreviations: CI, confidence interval; UFO, "U Find Out."
* $P < 0.01$.
[a] Time frame is past year for participants surveyed in 1997–1999.
[b] Data not collected in wave 1. Time frame is past week for participants surveyed in 2003–2007.

Table 2. Age-, Race-, and Sex-adjusted Standardized Mortality Ratios for Young Injection Drug Users in San Francisco, California, Using National Reference, 1999–2007

Year	Overall				Females			Males		
	UFO Mortality Rate/1,000 Person-Years	Crude National Mortality Rate/1,000 Person-Years	Adjusted SMR	95% CI	Mortality Rate/1,000 Person-Years	Adjusted SMR	95% CI	Mortality Rate/1,000 Person-Years	Adjusted SMR	95% CI
1999–2001	13.1	8.5	15.3	7.6, 27.3	7.1	20.0	2.4, 72.3	16.1	14.5	6.6, 27.6
2002–2004	9.4	8.4	10.6	5.8, 17.8	5.9	15.8	3.3, 46.1	11.3	9.7	4.9, 17.4
2005–2007	7.6	8.1	8.3	4.4, 14.3	7.2	19.1	5.2, 48.8	7.8	6.7	3.1, 12.7

Abbreviations: CI, confidence interval; SMR, standardized mortality ratio; UFO, "U Find Out."

Table 3. Incidence of Overdose Mortality Among Young Injection Drug Users in San Francisco, California, 1999–2007

Year	Overall		Females		Males	
	Mortality Rate/1,000 Person-Years	95% CI	Mortality Rate/1,000 Person-Years	95% CI	Mortality Rate/1,000 Person-Years	95% CI
1999–2001	7.1	3.2, 15.9	0		10.7	4.8, 23.8
2002–2004	6.0	3.1, 11.6	2.0	0.3, 13.9	8.2	4.1, 16.4
2005–2007	4.1	2.0, 8.6	1.8	0.3, 12.7	5.2	2.3, 11.6

Abbreviation: CI, confidence interval.

95% CI: 0.88, 3.15) at $P < 0.05$. Mortality incidence did not differ by age, race/ethnicity, ever having overdosed, duration of injection, or incarceration in the past 3 months.

By use of national mortality as the reference, the overall adjusted standardized mortality ratio for 1999–2001 was 15.3 (95% CI: 7.6, 27.3), which decreased to 10.6 (95% CI: 5.8, 17.8) in 2002–2004 and 8.3 (95% CI: 4.4, 14.3) in 2005–2007 (Table 2). When data were stratified by wave of recruitment, temporal declines in mortality persisted (data not shown). In females, the adjusted standardized mortality ratios were higher than in males across all time bands. Furthermore, there did not appear to be a trend toward a decreasing standardized mortality ratio over time among women, as there appeared to be for men. Incidence of overdose mortality also declined between 2000 and 2007 (Table 3). However, women had lower rates of overdose mortality compared with men, and their rates did not appear to decline. The principal cause of death for the overall cohort was overdose (57.9%), followed by self-inflicted injuries (13.2%), other drug-related medical conditions (13.2%), trauma or accidents (10.5%), and other causes (5.3%).

In Cox regression with time-varying covariates adjusting for age and duration injecting, participants who used heroin or heroin mix most days or who had a recent overdose had elevated mortality (Table 4). Of these factors, heroin/heroin mix as the drug used most days in the past month was the only one independently associated with mortality (adjusted hazard ratio = 5.76, 95% CI: 1.37, 24.30) ($P < 0.05$). Having a recent nonfatal overdose was associated with higher relative hazards of death, although this did not quite meet statistical significance. After adjustment for other covariates, being HCV infected was not significantly associated with increased relative hazards of death.

DISCUSSION

This study of young IDUs in San Francisco between 1997 and 2007 found overall mortality rates 10 times higher than those in the general population. Mortality appeared to decline over the 10-year period; however, stratified results suggested that declines were restricted to males. The leading cause of death in this cohort was overdose, and primary use of heroin was the only significant independent risk factor for death that was observed in the study. Together, these findings highlight the substantial mortality risk associated with injecting drugs for young persons and point to a continued need for public health interventions that address the risk of overdose in this population in order to reduce premature deaths.

The overall mortality rate observed in this study of IDUs in San Francisco was 9.12 per 1,000 person-years of observation. Another study of young and recent-onset injectors that included data from 5 different cities in the United States

Table 4. Multivariate Cox Proportional Hazards Model of Independent Predictors of Mortality Among Young Injection Drug Users in San Francisco, California, 1997–2007[a]

Characteristic	Adjusted Hazard Ratio	95% CI	*P* Value
Study wave			
1997–1999	1.00	Referent	
2000–2002	0.83	0.39, 1.75	0.62
2003–2007	0.82	0.28, 2.42	0.72
Age, years	1.00	0.89, 1.11	0.98
Duration injecting, years			
≤2	1.00	Referent	
>2–5	0.94	0.35, 2.51	0.90
>5	1.50	0.42, 5.36	0.54
Gender			
Female	1.00	Referent	
Male	1.62	0.73, 3.58	0.24
Hepatitis C virus status			
Negative	1.00	Referent	
Positive	1.36	0.68, 2.70	0.39
Drug injected most days in past month			
Methamphetamine	1.00	Referent	
Heroin/heroin mix	5.76	1.37, 24.30	0.02
Other	3.52	0.31, 39.44	0.31
Overdosed, past 3 months[b]			
No	1.00	Referent	
Yes	1.92	0.91, 4.06	0.09

Abbreviation: CI, confidence interval.

[a] Includes all variables associated at $P < 0.10$ in bivariate analyses and adjusting for age, duration injecting, gender, hepatitis C virus status, and study wave.

[b] Time frame is past year for participants surveyed in 1997–1999.

found a similar mortality rate (3). That study found an initial increase of mortality in the 2 years of the cohort study, followed by a decrease, which led the authors to hypothesize that the immediate years following initiation of drug use might be a period of higher risk. In contrast, our study found only a continuous decline in mortality over calendar periods. The prior study by Vlahov et al. enrolled only recently initiated IDUs (<5 years) and therefore did not include a range of age and duration of drug use to clarify period effects. Our study included participants with a wide range of injecting exposure time (median = 4 years, IQR: 1.7–6.7), and we performed multivariate Cox models to analyze the independent effects of age and duration of use. The results did not support an association between duration of use and mortality.

There are several potential explanations for the observed decline in mortality in our cohort of young injectors. First, it is possible that public health interventions could have led to fewer deaths over the years. In the past decade, several overdose prevention programs have been implemented in the United States including the distribution of the overdose-reversal drug, opiate antagonist naloxone (NARCAN; Endo Pharmaceuticals, Inc., Chadds Ford, Pennsylvania), to IDUs (12–16). One of the longest running naloxone prescription programs in the United States was implemented in San Francisco in 2003 and has provided overdose prevention training and prescribed naloxone to nearly 2,000 individuals as of December 2009 (16). Furthermore, improved access to treatment for opioid dependence may be a factor, as opioid substitution treatment in IDUs has been shown to be a significant factor in reducing the risk of mortality (17). There could have been a change in the purity of heroin over the years (18). Finally, because there was ongoing enrollment of our cohort, differences in mortality over time may be associated with unmeasured differences in the UFO Study population over time, such that subjects with lower risk were recruited during later phases of the study. To test this hypothesis, we included a variable for period of study recruitment in our Cox models. Although we did not find significant differences in the hazard ratios for death by enrollment wave, the hazard ratios were less than 1.0 comparing the later with earlier years, suggesting that this could be an underlying factor. Furthermore, when we investigated changes in the UFO Study population across the 3 waves of data collection, we found that participants in the first wave were more likely to inject heroin or heroin mixed with other drugs in the past month and were significantly less likely to inject methamphetamine in the past month at baseline (data not shown). Since primary use of heroin as the drug of injection was strongly predictive of death in our Cox models (hazard ratio = 5.76, 95% CI: 1.37, 24.30), this observed decline in heroin injection in later waves of study recruitment is a likely contributor to the decline in overdose mortality over time.

Our study of young injection drug users found the most common causes of death to be overdose and trauma, which is consistent with prior studies (3, 6, 7, 19). This supports a continued focus on overdose prevention in this age group. Of the risk factors examined in this study, only use of heroin as the main injecting drug was independently predictive of death. This contrasts with 2 prior studies of young IDUs that found HIV infection to be significantly associated with mortality (3, 6). This difference can likely be explained by the relatively small absolute number of participants included in our sample and the small number of persons with HIV. Also, HIV as a cause of death has been shown to have peaked in the 1990s (20, 21), before the advent of highly active antiretroviral therapy, and is less likely to impact mortality in young versus older injectors (7).

Our study showed differences in death rates between men and women. Although the incidence of death was higher among men compared with women, standardized mortality ratios were higher for women. Female injectors were 12 times more likely to die compared with females in the general US population; in contrast, male injectors were 8 times more likely to die compared with the male US population. A similar difference in sex-stratified standardized mortality ratios has been reported in a cohort of young IDUs in Canada (6). In addition, although rates of death appeared to decline over time among men, rates appeared relatively stable for women over the 10-year period. Sex-stratified results should be interpreted with caution however, as the sample size was relatively small and resulted in wide confidence intervals for period-specific mortality rates.

There were a number of important limitations to this study. Mortality rates may be underestimates because of deaths that were not captured by the NDI search criteria. A number of study participants were believed to have given pseudonyms rather than birth names, which would not allow matching, and incorrect birth dates may have been given. Causes of death were based on ICD codes on death certificates that may be inaccurate (22). Analyses used baseline predictor data, and some covariates measured constructs that could change over time (HIV and HCV infections, injecting behaviors, incarceration, and so on). Therefore, analyses of certain predictors should be interpreted with caution. Finally, the relatively modest sample size and small absolute number of deaths are limitations, particularly with regard to sex-stratified results.

In conclusion, this study of young IDUs in San Francisco demonstrated that injection drug use is associated with substantially increased risk for death, and that most of the excess risk is due to overdose and trauma. Results also suggest that, among young adult injectors in San Francisco, mortality rates are decreasing over time but are generally restricted to males. Research is needed to confirm these results and explore reasons for sex-specific differences in mortality in young IDUs.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California (Jennifer L. Evans, Kimberly Page); Department of Medicine, University of California, San Francisco, San Francisco, California (Judith A. Hahn, Paula J. Lum); School of Medicine, Boston University, Boston, Massachusetts (Judith I. Tsui); and Department of Medicine, University of California, San Diego, San Diego, California (Peter J. Davidson).

This research was conducted with the support of a grant from the National Institute on Drug Abuse (5R01DA016017-08).

The authors would like to thank the UFO Study Field Staff for their time and dedication in conducting the study.

Conflict of interest: none declared.

REFERENCES

1. Degenhardt L, Hall W, Warner-Smith M. Using cohort studies to estimate mortality among injecting drug users that is not attributable to AIDS. *Sex Transm Infect*. 2006;82(suppl 3): iii56–iii63.
2. Hulse GK, English DR, Milne E, et al. The quantification of mortality resulting from the regular use of illicit opiates. *Addiction*. 1999;94(2):221–229.
3. Vlahov D, Wang C, Ompad D, et al. Mortality risk among recent-onset injection drug users in five U.S. cities. Collaborative Injection Drug User Study. *Subst Use Misuse*. 2008; 43(3-4):413–428.
4. Bargagli AM, Hickman M, Davoli M, et al. Drug-related mortality and its impact on adult mortality in eight European countries. COSMO European Group. *Eur J Public Health*. 2006;16(2):198–202.
5. Degenhardt L, Bucello C, Mathers B, et al. Mortality among regular or dependent users of heroin and other opioids: a systematic review and meta-analysis of cohort studies. *Addiction*. 2011;106(1):32–51.
6. Miller CL, Kerr T, Strathdee SA, et al. Factors associated with premature mortality among young injection drug users in Vancouver. *Harm Reduct J*. 2007;4:1. (doi:10.1186/1477-7517-4-1).
7. Clausen T, Waal H, Thoresen M, et al. Mortality among opiate users: opioid maintenance therapy, age and causes of death. *Addiction*. 2009;104(8):1356–1362.
8. Gossop M, Stewart D, Treacy S, et al. A prospective study of mortality among drug misusers during a 4-year period after seeking treatment. *Addiction*. 2002;97(1):39–47.
9. Hahn JA, Page-Shafer K, Lum PJ, et al. Hepatitis C virus seroconversion among young injection drug users: relationships and risks. *J Infect Dis*. 2002;186(11):1558–1564.
10. Lum PJ, Ochoa KC, Hahn JA, et al. Hepatitis B virus immunization among young injection drug users in San Francisco, Calif: the UFO Study. *Am J Public Health*. 2003; 93(6):919–923.
11. Page K, Hahn JA, Evans J, et al. Acute hepatitis C virus infection in young adult injection drug users: a prospective study of incident infection, resolution, and reinfection. *J Infect Dis*. 2009;200(8):1216–1226.
12. Seal KH, Thawley R, Gee L, et al. Naloxone distribution and cardiopulmonary resuscitation training for injection drug users to prevent heroin overdose death: a pilot intervention study. *J Urban Health*. 2005;82(2):303–311.
13. Baca CT, Grant KJ. Take-home naloxone to reduce heroin death. *Addiction*. 2005;100(12):1823–1831.
14. Maxwell S, Bigg D, Stanczykiewicz K, et al. Prescribing naloxone to actively injecting heroin users: a program to reduce heroin overdose deaths. *J Addict Dis*. 2006;25(3):89–96.
15. Sherman SG, Gann DS, Tobin KE, et al. ''The life they save may be mine'': diffusion of overdose prevention information from a city sponsored programme. *Int J Drug Policy*. 2009; 20(2):137–142.
16. Enteen L, Bauer J, McLean R, et al. Overdose prevention and naloxone prescription for opioid users in San Francisco. *J Urban Health*. 2010;87(6):931–941.
17. Kimber J, Copeland L, Hickman M, et al. Survival and cessation in injecting drug users: prospective observational study of outcomes and effect of opiate substitution treatment. *BMJ*. 2010;340:c3172. (doi:10.1136/bmj.c3172).
18. Cunningham JK, Maxwell JC, Campollo O, et al. Proximity to the US-Mexico border: a key to explaining geographic variation in US methamphetamine, cocaine and heroin purity. *Addiction*. 2010;105(10):1785–1798.
19. Webb L, Oyefeso A, Schifano F, et al. Cause and manner of death in drug-related fatality: an analysis of drug-related deaths recorded by coroners in England and Wales in 2000. *Drug Alcohol Depend*. 2003;72(1):67–74.
20. Copeland L, Budd J, Robertson JR, et al. Changing patterns in causes of death in a cohort of injecting drug users, 1980–2001. *Arch Intern Med*. 2004;164(11):1214–1220.
21. Vlahov D, Wang CL, Galai N, et al. Mortality risk among new onset injection drug users. *Addiction*. 2004;99(8):946–954.
22. Nielsen GP, Björnsson J, Jonasson JG. The accuracy of death certificates. Implications for health statistics. *Virchows Arch A Pathol Anat Histopathol*. 1991;419(2):143–146.

Factors associated with difficult electronic health record implementation in office practice

Marshall Fleurant,[1] Rachel Kell,[2] Chelsea Jenter,[3] Lynn A Volk,[4] Fang Zhang,[5,6] David W Bates,[3,4,7,8] Steven R Simon[3,7,9]

[1]Boston Medical Center, Boston, Massachusetts, USA
[2]Massachusetts eHealth Collaborative, Waltham, Massachusetts, USA
[3]Division of General Internal Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA
[4]Partners HealthCare System, Inc., Wellesley, Massachusetts, USA
[5]Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA
[6]Department of Population Medicine, Boston, Massachusetts, USA
[7]Harvard Medical School, Boston, Massachusetts, USA
[8]Harvard School of Public Health, Boston, Massachusetts, USA
[9]Section of General Internal Medicine, VA Boston Healthcare System, Boston, Massachusetts, USA

Correspondence to
Dr Marshall Fleurant, Boston Medical Center, Boston University School of Medicine, Section of General Internal Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA;
marshall.fleurant@bmc.org

Received 6 November 2011
Accepted 14 December 2011
Published Online First
16 January 2012

ABSTRACT

Little is known about physicians' perception of the ease or difficulty of implementing electronic health records (EHR). This study identified factors related to the perceived difficulty of implementing EHR. 163 physicians completed surveys before and after the implementation of EHR in an externally funded pilot program in three Massachusetts communities. Ordinal hierarchical logistic regression was used to identify baseline factors that correlated with physicians' report of difficulty with EHR implementation. Compared with physicians with ownership stake in their practices, physician employees were less likely to describe EHR implementation as difficult (adjusted OR 0.5, 95% CI 0.3 to 1.0). Physicians who perceived their staff to be innovative were also less likely to view EHR implementation as difficult (adjusted OR 0.4, 95% CI 0.2 to 0.8). Physicians who own their practice may need more external support for EHR implementation than those who do not. Innovative clinical support staff may ease the EHR implementation process and contribute to its success.

Physicians who have not demonstrated 'meaningful use' of health information technology (HIT) by 2015 may face reduced future Medicare reimbursement.[1] Failure to implement electronic health records (EHR) may prove costly to individual physicians and may make it hard for them to engage in care redesign that will be needed under healthcare reform.[2 3] Awards up to US $44 000 through Medicare or even US$63 750 through Medicaid are currently available to physicians in office practice who demonstrate meaningful use of their EHR over 5 years.[1] To receive the maximum reimbursement providers must begin participating by 2012.

Through the federal government, health information technology regional extension centers (REC) are positioned across the country to offer technical support and best practices to at least 100 000 providers over 2 years.[4] In addition, approximately US$500 million has been awarded in federal support to states and territories to increase health information exchange.[5 6]

Earlier studies have identified factors that predispose practices to successful EHR adoption, namely larger practices, those affiliated with large hospitals, or those involved with teaching.[7–10] Physicians face numerous barriers to participating in the meaningful use program, including a bewildering number of choices when selecting an EHR.[11] It is unclear how fast providers will adopt. Ford *et al*,[12] using EHR adoption data from six previous surveys, found under their most optimistic scenario a predicted EHR adoption rate of 61% by 2014, far short of 'widespread adoption'.

In contrast to adoption, comparatively little is known about the factors that predict successful implementation, the installation of EHR and their incorporation into the usual operations of a functioning practice. Few data are available regarding what factors influence physicians' perception of the ease or difficulty of EHR implementation. This perception of ease or difficulty is important. It may influence the actual successfulness of EHR implementation or even the decision to complete an EHR implementation that is underway. Therefore, we undertook the present study, using data from practices in three communities participating in the Massachusetts eHealth Collaborative (MAeHC) EHR implementation project. Our objective was to identify factors present at baseline that were subsequently related to the perceived difficulty of implementing EHR.

METHODS

Study design

We carried out a pre–post evaluation of the MAeHC three-community intervention to promote physicians' adoption and implementation of EHR. We collected physician attitudes, demographic and practice characteristics and used post-intervention surveys to measure physicians' perceived difficulty of EHR implementation. The Partners HealthCare Human Research Committee approved the study protocol.

Intervention

The details of the MAeHC program have been described previously.[13–15] In summary, between 2006 and 2008, MAeHC installed robust EHR in physician practices in three Massachusetts communities. MAeHC provided extensive on-site consultation, at no cost to these practices, to facilitate workflow redesign and integration of the EHR into the practice. They also provided technical support related to hardware and software.

Settings and participants

The Massachusetts communities of Brockton, Newburyport and North Adams were selected for the MAeHC program through a competitive process described in detail in Goroll *et al*.[13] Briefly, MAeHC issued a request for applications to which 35 communities throughout Massachusetts submitted applications to participate. MAeHC selection committee members made site visits to

six finalist communities and ultimately selected the three participant communities. Selection criteria emphasized local physician leadership and community support of the application.

A total of 167 physician practices, representing 91% of all eligible primary care and specialty practices, participated. We identified all physicians in each of the participating practices for pre-intervention (2005) and post-intervention (2009) surveys. In the pre-intervention survey we identified 500 physicians from 167 practices; 355 completed the survey (response rate 77%).[16] In 2009, 468 physicians were eligible, of whom 319 completed the survey (response rate 68%). For this study we included only the 163 physicians from 134 practice sites who completed both the 2005 and 2009 survey questionnaires.

Survey design

The design and content of the pre and post-intervention surveys were based on similar statewide surveys of physicians described previously.[16 17] Briefly, the 2005 survey included items intended to measure physicians' attitudes toward the use of computers in healthcare, perceptions of the quality of care, physician demographics and practice characteristics. The post-intervention (2009) survey retained verbatim all of the items from the original survey and incorporated new items that would not have been relevant before the intervention. For example, new questions asked about health information exchange, the EHR implementation process and workflow analysis.

Survey administration and data collection

Survey administration and data collection methods in 2005 and 2009 were similar. In 2005, paper surveys were hand-delivered to physicians' offices by MAeHC practice consultants, who had been making routine visits to the practices in preparation for the anticipated implementation. Physicians had the option of returning the survey directly to practice consultants or by mail.[16] In 2009, the Survey Lab of the University of Chicago, a third-party research and consulting service, administered the survey to MAeHC physicians; these physicians returned the completed survey by mail or via the web to the Survey Lab in Chicago.

Main outcome measure

Our main outcome was the response to the following question in the post-intervention survey: 'Was the implementation process for your EHR… (a) very difficult; (b) somewhat difficult; (c) not difficult?'

Other variables

The pre-intervention survey included physician characteristics (age, sex, race, years in practice, number of outpatient visits per week); practice characteristics (number of physicians in practice, specialty, ownership, and financial resources available for expansion); and information on individual physician stressors (isolation from colleagues, working long hours, personal/job stress, and demoralization). Several dimensions of organizational culture were assessed, including innovation, initiative, quality improvement, evaluation, and risk assessment (see table 1). We requested information on the availability and physicians' use of 10 key functions of the EHR post-intervention to calculate an EHR usage score.[9] As in previous analyses, we calculated an EHR usage score as the number of functions reported as used most or all of the time, divided by the number of available functions. For example, physicians who reported using none of their available EHR functions would score 0, while those who reported using all of their available functions most or all of the time would score 1. In another example, physicians who reported using four of their eight available functions most or all of the time would score 0.5.

Table 1 Characteristics of physicians and practices

Characteristic	N (%)
Total respondents	156 (100%)
Age, years	
≤50	89 (57%)
>50	67 (43%)
Male	119 (76%)
White	126 (81%)
Years with current practice group	
<10	85 (54%)
10–19	39 (25%)
≥20	32 (21%)
Practice owner, N (%)*	104 (67%)
Primary care	58 (37%)
Practice size	
1–2 physicians	39 (25%)
3–5 physicians	26 (17%)
>5 physicians	91 (58%)
Outpatients visit per week†	
<50	38 (27%)
50–100	60 (43%)
>100	42 (30%)
Resources available for practice expansion‡	
Extensive or moderate resources	28 (18%)
Limited or no resources	105 (68%)
Physician stressors, considering it a problem, N (%)§	
Feeling demoralized	69 (44%)
Having to work long hours	121 (78%)
Personal or professional stress	123 (79%)
Isolation from colleagues	132 (84%)
Organizational culture¶	
The office staff are innovative	67 (43%)
Physician(s) are innovative	50 (32%)
First to discover new treatments**	105 (67%)
Improving quality of care††	20 (13%)
Evaluate quality changes‡‡	53 (34%)
We have quality problems	122 (78%)
We are good at preventing errors§§	87 (58%)

Percentages do not necessarily add to 100% because of rounding, missing data not included in percentages.
*Practice owner defined as physicians with full or partial ownership stake in their practice.
†Missing 16 responses.
‡Refers to capital available for future growth by practice; missing 23 responses.
§Proportion of physicians who view each category as a problem versus not a problem.
¶Proportion of physicians who agree that each category represents their office environment.
**Among my colleagues, I am usually the first to discover about new treatments and diagnostic tests.
††We are actively doing things to improve quality of care.
‡‡After we make changes to improve quality, we evaluate their effectiveness.
§§Our procedures and systems are good at preventing errors.

Statistical analysis

To identify potential confounders between our main outcome and additional covariates we calculated χ^2 tests or Fisher's exact test among categorical values and Wilcoxon rank-sum tests for continuous variables. Variables noted to be significantly associated ($p<0.05$) with the main outcome measure were included in the multivariate model. For multivariate analysis, we performed ordinal hierarchical logistic regression to account for the clustering of physicians within practice; the ordinal logistic regression assumption was met. We decided a priori to include sex and categorical age in the multivariate model. We validated use of

the ordinal model using the score test for proportional odds assumption.

RESULTS

Overall, 163 physicians completed both surveys in 2005 and 2009. Among these physicians, 156 physicians (96%) answered questions relating to our main outcome. The physicians were mostly men, and most had full or partial ownership stake in their practice (see table 1). Most participants (70%) had been in practice for at least 15 years, with a median of 22 years since graduation from medical school. A total of 22 physicians (15%) reported having a 'solo' practice; 99 physicians (63%) indicated that they practised in a primary care or single specialty group practice and 35 (22%) worked in multispecialty group practices. A total of 26 physicians (17%) indicated that their practice had at least moderate resources for expansion or improvements of any kind; only two physicians (1%) stated that their practice had extensive resources for expansion. The mean EHR usage score was 0.68 (SD=0.28).

Overall, 54 physicians (35%) reported that the implementation process for their EHR was very difficult. Eighty-four (54%) of the physicians found the process to be somewhat difficult; only 18 (12%) reported their implementation process was not difficult.

Factors associated with difficulty of EHR implementation in bivariate analyses are shown in table 2. Compared with physicians who were partial or full owners of their practice, those without any ownership stake in their practice were less likely to view the implementation process as difficult or very difficult (OR 0.5, 95% CI 0.2 to 0.9). In terms of absolute numbers, 26% of non-owners (14/54) indicated that implementation was very difficult, compared with 38% of owners (40/104).

Physicians who indicated that the office staff was innovative were less likely to view the implementation process as difficult or very difficult (OR 0.4, 95% CI 0.2 to 0.8), compared with physicians who did not indicate that their staff was innovative. Higher EHR usage scores were associated with lower likelihood of viewing the implementation process as difficult or very difficult (OR 0.3, 95% CI 0.1 to 0.9).

In multivariate analysis, however, physicians' EHR usage score was not associated with the perceived difficulty of EHR implementation. Physicians who were employed by their practice (ie, those who did not have partial or full ownership stake in the practice) were less likely to view EHR implementation as difficult (adjusted OR 0.5, 95% CI 0.3 to 1.0). Similarly, physicians who perceived their staff as innovative were less likely to view EHR implementation as difficult (adjusted OR 0.4, 95% CI 0.2 to 0.8). We did not find evidence of effect modification by practice size on owners' perception of difficulty (p=0.69).

Table 2 The odds of finding the EHR implementation process difficult by physicians in ambulatory practices*

	Unadjusted OR (95% CI)	Adjusted OR (95% CI)†
Male	1.0 (0.5 to 2.0)	0.7 (0.3 to 1.6)
Age >50 years	1.4 (0.8 to 2.5)	1.3 (0.7 to 2.5)
EHR usage score‡	**0.3 (0.1 to 0.9)**	0.3 (0.1 to 1.1)
Non-owners	**0.5 (0.2 to 0.9)**	**0.5 (0.3 to 1.0)**
Staff innovation	**0.4 (0.2 to 0.8)**	**0.4 (0.2 to 0.8)**

*Bold items indicate a statistically significant value at p≤0.05.
†Adjusted for sex, categorical age, calculated electronic health record (EHR) usage score, ownership, proportion of physicians who agree with the statement that their office staff are innovative.
‡Measured as a continuous variable.

DISCUSSION

We evaluated how physicians practising in three communities in Massachusetts perceived the EHR implementation process and factors that facilitated EHR implementation. In this setting, we found that physicians who were not owners of the practice and those who perceived their office staff to be innovative were less likely to view the implementation process as difficult or very difficult.

Earlier studies have demonstrated that physicians working in larger practices and those affiliated with hospitals are more likely to have adopted EHR and to use those systems meaningfully.[7 9 10] However, relatively little attention has been paid to the ease or difficulty physicians perceive when they implement EHR in their practices.

We found that physicians who were owners of their practice perceived their EHR implementation to be more difficult, compared with physicians who did not have an ownership stake. Because ownership is generally associated with greater levels of responsibility for day-to-day practice operations and management, these physicians probably experienced more underlying challenges associated with EHR implementation and workflow transformation. Physician owners probably bear financial risk for failure of implementation out of proportion to payers or publicly funded health plans who benefit from patient safety and quality but not as directly from practice efficiency or revenue cycle management.[18] These physician owners may need not only financial support but also training and expert consultation to bolster the implementation process.

The association between physicians' perception of having innovative staff and their report of less difficult EHR implementation deserves exploration. Having staff who are innovative may reflect an office culture that is committed to progress or potentially a marker for a physician who values the role of non-physician staff in ensuring the smooth operations of the office. Understanding the role of 'innovative staff' may uncover a potential facilitator for easing EHR adoption and implementation.

This study has several limitations. Although the sample includes a broad range of primary care physicians and specialists from three diverse communities across Massachusetts, the findings may not generalize to practices in other settings. It should be noted that the physicians in this study were part of a well-supported EHR implementation program. While the HITECH-mandated REC program is supporting practices, it is likely that most practices nationally will receive less support from their REC than the physician groups in our study. We also note the limitation as with any survey-based research, that our variables are self-reported perceptions, rather than objective measures, and subject to a variety of biases, such as social desirability. Nonetheless, these perceptions are important, because they can influence physicians' behavior.

CONCLUSIONS

Physicians who have ownership stake in their practices experience first-hand the challenges of EHR implementation, and they generally perceive EHR implementation as being more difficult than their colleagues who are employed by the practices in which they work. Physicians who report that their staff is innovative perceive less difficulty with EHR implementation, suggesting that office staff play a critical role in EHR implementation, as they do in any transformation effort. Efforts to expand EHR implementation should focus on the needs of physician owners, and future efforts should emphasize the important role that non-physician staff members play in the process of innovation and transformation.

Acknowledgments The authors would like to thank the physicians who participated in the MAeHC for completing the surveys. The authors also thank the anonymous peer reviewer who provided helpful suggestions to improve the manuscript.

Funding This study was funded in part by the Agency for Healthcare Research and Quality cooperative agreement no 1UC1HS015397 and the Massachusetts e-Health Collaborative. MF was supported by a Ruth L. Kirschstein National Research Service Award grant no T32 HP12706.

Competing interests None.

Ethics approval The Partners HealthCare Human Research Committee approved the study protocol.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **US Department of Health and Human Services.** *EHR Incentive Program. Overview*. Centers for Medicare and Medicaid Services, 2011. http://www.cms.gov/ehrincentiveprograms/ (accessed 27 Jul 2011).
2. **Rudin RS,** Simon SR, Volk LA, *et al*. Understanding the decisions and values of stakeholders in health information exchanges: experiences from Massachusetts. *Am J Public Health* 2009;**99**:950—5.
3. **Adler-Milstein J,** Bates DW, Jha AK. A survey of health information exchange organizations in the United States: implications for meaningful use. *Ann Intern Med* 2011;**154**:666—71.
4. **US Department of Health and Human Services.** *The Office of the National Coordinator for Health Information Technology. Regional Extension Centers*. 2011. http://healthit.hhs.gov/portal/server.pt?open=512&objID=1495&parentname=CommunityPage&parentid=58&mode=2&in_hi_userid=11113&cached=true (accessed 27 Jul 2011).
5. **US Department of Health and Human Services.** *The Office of the National Coordinator for Health Information Technology. Beacon Community Program*. 2011. http://healthit.hhs.gov/portal/server.pt?open=512&objID=1805&parentname=CommunityPage&parentid=2&mode=2&cached=true (accessed 27 Jul 2011).
6. **US Department of Health and Human Services.** *The office of the National Coordinator for Health Information Technology. Health Information Exchange Challenge Grant Program*. 2011. http://healthit.hhs.gov/portal/server.pt?open=512&mode=2&objID=3378 (accessed 27 Jul 2011).
7. **Jha AK,** DesRoches CM, Campbell EG, *et al*. Use of electronic health records in U.S. hospitals. *N Engl J Med* 2009;**360**:1628—38.
8. **DesRoches CM,** Campbell EG, Rao SR, *et al*. Electronic health records in ambulatory care—a national survey of physicians. *N Engl J Med* 2008;**359**:50—60.
9. **Simon SR,** Kaushal R, Cleary PD, *et al*. Physicians and electronic health records: a statewide survey. *Arch Intern Med* 2007;**167**:507—12.
10. **Simon SR,** Kaushal R, Cleary PD, *et al*. Correlates of electronic health record adoption in office practices: a statewide survey. *J Am Med Inform Assoc* 2007;**14**:110—17.
11. **Balfour DC,** Evans S, Januska J, *et al*. Health information technology—results from a roundtable discussion. *J Manag Care Pharm* 2009;**15**(1 Suppl A): 10—17.
12. **Ford EW,** Menachemi N, Phillips MT. Predicting the adoption of electronic health records by physicians: when will health care be paperless? *J Am Med Inform Assoc* 2006;**13**:106—12.
13. **Goroll AH,** Simon SR, Tripathi M, *et al*. Community-wide implementation of health information technology: the Massachusetts eHealth Collaborative experience. *J Am Med Inform Assoc* 2009;**16**:132—9.
14. **Halamka J,** Aranow M, Ascenzo C, *et al*. Health care IT collaboration in Massachusetts: the experience of creating regional connectivity. *J Am Med Inform Assoc* 2005;**12**:596—601.
15. **Mostashari F,** Tripathi M, Kendall M. A tale of two large community electronic health record extension projects. *Health Aff (Millwood)* 2009;**28**:345—56.
16. **Simon SR,** Kaushal R, Jenter CA, *et al*. Readiness for electronic health records: comparison of characteristics of practices in a collaborative with the remainder of Massachusetts. *Inform Prim Care* 2008;**16**:129—37.
17. **Quinn MA,** Wilcox A, Orav EJ, *et al*. The relationship between perceived practice quality and quality improvement activities and physician practice dissatisfaction, professional isolation, and work-life stress. *Med Care* 2009;**47**:924—8.
18. **Middleton B,** Hammond WE, Brennan PF, *et al*. Accelerating U.S. EHR adoption: how to get there from here. Recommendations based on the 2004 ACMI retreat. *J Am Med Inform Assoc* 2005;**12**:13—19.

Women & Health, 52:454–471, 2012

ISSN: 0363-0242 print/1541-0331 online
DOI: 10.1080/03630242.2012.684088

Routledge
Taylor & Francis Group

Posttraumatic Stress Disorder and Intimate Partner Violence in a Women's Headache Center

MEGAN R. GERBER, MD, MPH
VA Boston Healthcare System/Women's Health, Boston University School of Medicine, Boston, Massachusetts, USA

LISE E. FRIED, DSc, MS
Institute for Community Health, Harvard Medical School, Cambridge, Massachusetts, USA

SUZANNE L. PINELES, PhD and JILLIAN C. SHIPHERD, PhD
VA Boston Healthcare System, National Center for PTSD/Women's Health Sciences Division, Boston University School of Medicine, Boston, Massachusetts, USA

CAROLYN A. BERNSTEIN, MD
Beth Israel Deaconess Medical Center, Harvard Medical Faculty Physicians Comprehensive Headache Center, Harvard Medical School, Boston, Massachusetts, USA

Background: Posttraumatic stress disorder has been linked to women's ill health, including headaches. Intimate partner violence, which may result in posttraumatic stress disorder, is often reported by women with headaches. Prior studies of intimate partner violence and headache have estimated lifetime but not 12-month prevalence. The researchers in this study examined the relationship between headache and posttraumatic stress disorder in a novel population, and estimated 12-month and lifetime prevalence rates of intimate partner violence. Methods: Patients were recruited from a women's headache center (n = 92) during 2006–07 and

Received September 20, 2011; revised April 3, 2012; accepted April 4, 2012.

Funding Source: Institute for Community Health, Cambridge, Massachusetts, intramural funding. The authors disclose no commercial associations that might create a conflict of interest.

The authors thank Connie Marchurs, Jessica Barrett, MPH, Amanda Robinson, MD, Institute for Community Health (Cambridge, MA), and Margaret Bauer, BA.

Address correspondence to Megan R. Gerber, MD, MPH, VA Boston Healthcare System/Women's Health, Boston University School of Medicine, 150 S. Huntington Ave., Jamaica Plain, MA 02130. E-mail: megan.gerber@va.gov

completed the Migraine Disability Assessment measure of headache severity. Posttraumatic stress disorder was measured using a modified Breslau scale. Twelve-month and lifetime physical intimate partner violence were measured with the Partner Violence Screen and the STaT ("slapped, threatened and throw") measure. Multivariable regression determined factors independently associated with headache severity. Results: Among all participants, 28.3% screened positive for posttraumatic stress disorder; 9.8% and 36.9% of women endorsed recent and lifetime intimate partner violence. Posttraumatic stress disorder was strongly associated with headache severity ($\beta = 34.12$, $p = 0.01$). Patients reporting lifetime intimate partner violence exhibited a trend of nine additional days of disability due to headache over 90 days. Conclusions: Posttraumatic stress disorder and intimate partner violence occur among a sizable proportion of women referred for headache. The authors' findings reaffirm that clinicians treating women with headaches must be aware of the possibility of posttraumatic stress disorder and intimate partner violence in such patients.

KEYWORDS posttraumatic stress disorder, intimate partner violence, headache, chronic pain, disability

INTRODUCTION

Interpersonal trauma has a profound and lasting impact on women's long-term health (Schnurr & Green, 2004). Many forms of chronic pain have been linked to past traumatic experiences (Paras et al., 2009; Romans et al., 2002; Wuest et al., 2008; Walling et al., 1994). Headache is one of the most frequent forms of chronic pain, accounting for considerable disability and use of health services (Lipton et al., 2007), with a disproportionate impact on women (Diamond, 2007). Over the last decade, an enlarging body of work has examined the association between headache and exposure to many forms of interpersonal trauma (Anda et al., 2010; Tietjen et al., 2009a, 2009b; Golding, 1999). Emerging data has suggested that posttraumatic stress disorder (PTSD) is an important contributory factor in the relationship between trauma and resultant physical health sequelae (Schnurr & Green, 2004; Kimerling, Clum, & Wolfe, 2000; Dennis et al., 2009; Wuest et al., 2009). PTSD is prevalent in ambulatory medical populations; recent studies have found a prevalence ranging from 12% to 33% among primary care patients (Liebschutz et al., 2007; Alim et al., 2006; Stein et al., 2000).

PTSD is characterized by re-experiencing, avoidance, numbing, and arousal symptoms related to a significant traumatic event (American

Psychological Association, 2000), and may be accompanied by altered allostatic load (physiologic changes that result from chronic stress) (Glover, Stuber, & Poland, 2006), which is one recognized pathway to poor health_ (Rasmusson et al., 2010; Dennis et al., 2009; Dutton et al., 2006; Gill, Szanton, & Page, 2005) and chronic pain (Peterlin et al., 2008; Raphael & Widom, 2010; Otis, Keane, & Kerns, 2003; Sharp & Harvey, 2001). PTSD has a higher incidence among women (Tolin & Foa, 2006). Although PTSD appears to play a key role in the relationship between trauma and subsequent poor physical health and chronic pain (Kimerling, Clum, & Wolfe, 2000; Sharp & Harvey, 2001; Schnurr & Jankowski, 1999), it is frequently misdiagnosed or under-recognized outside of mental health treatment settings (Liebschutz et al., 2007). Thus, primary care providers caring for women may fail to recognize the presence of PTSD among their patients.

Headache and PTSD often co-occur (Peterlin, Nijjar, & Tietjen, 2011). A large general population study revealed that the 12-month odds of PTSD in women with migraine were increased nearly four-fold (3.82), and lifetime odds were increased three-fold compared to women not reporting headache (Peterlin, Rosso, et al., 2011). The published prevalence rates of PTSD among headache patients have ranged from 12.5% to 42.9% (de Leeuw, Schmidt, & Carlson, 2005; Peterlin et al., 2008, 2009; Peterlin, Nijjar, et al., 2011). In a multi-site study of primary care patients, participants with PTSD had a two-fold odds of reporting headache (Lowe et al., 2011). The presence of PTSD was also a risk factor for chronic (versus episodic) headache in the same clinic population (Peterlin et al., 2008, 2009) serving both genders with a majority of female participants. Estimates from an all-women's headache center are lacking in the literature.

One frequent form of interpersonal trauma resulting in PTSD is intimate partner violence (IPV; Dutton et al., 2006). While PTSD can result from multiple forms of interpersonal trauma, the authors focused on IPV in the current study due to the relative paucity of reports on IPV in clinical settings dedicated to treating headaches. IPV is a pattern of coercive behavior in which one person attempts to control another through threats or actual use of physical violence, sexual assault, and verbal or psychological abuse (Heron et al., 2003). The population-based National Intimate Partner and Sexual Violence Survey (NISVS) found a lifetime rate of any severe physical IPV of 24.3%, and a rate of recent IPV of 2.7% among women. When less severe forms of physical IPV were included, the prevalence rates were 30.3% and 3.6%, respectively (Black et al., 2011). IPV has well-known adverse health effects (Coker et al., 2002; Campbell et al., 2002). Acute physical injury often occurs, with head and neck injuries the most frequent subtype of musculoskeletal injury resulting from IPV (Tam et al., 2010). Abused women have 1.5 to 2 times higher risk of chronic pain, including headache (Coker et al., 2002) (see Figure 1).

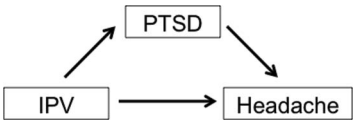


FIGURE 1 Relationship between headache severity, PTSD, and IPV.

Women who are victims of IPV are up to 60% more likely to suffer from headache than unabused controls (Coker et al., 2002; Cripe et al., 2011). A retrospective chart review of mostly female (90.1%) tertiary care headache patients in the United States found that nearly 35% reported a lifetime history of physical and/or sexual abuse (Peterlin et al., 2007). However, that study did not use standardized measures of abuse, making the results difficult to generalize. A recent study of post-partum hospitalized women in Peru found that 34.8% of migraine sufferers reported a lifetime history of IPV (Cripe et al., 2011). Neither of these studies ascertained recent or 12-month prevalence rates of IPV in the populations studied. Measurement of the prevalence of recent abuse would provide insight into the identification and management of women with headache who are recently and/or currently involved in abusive relationships. Intervention for women experiencing current or recent abuse differs from care needed by women with past history of IPV.

Given the growing popularity of outpatient, female-only clinics in the United States (American Hospital Association, 2000), including the expansion of dedicated primary care clinics for women veterans (Bean–Mayberry et al., 2003), examination of prevalence rates of IPV in ambulatory headache treatment settings that exclusively serve women would be useful. The majority of headache patients seen in primary care settings are not referred to specialists (Stokes et al., 2011). Therefore, describing the prevalence of potential PTSD and IPV in this setting could yield findings relevant to more severe headache patients who can be more challenging to manage in primary care.

Survivors of IPV and other forms of trauma may prefer women-only clinics because such patients are often hyper-vigilant and likely to overestimate the threat associated with being around men (Elwood et al., 2007). In previous studies of satisfaction with women-only practice settings versus traditional models, only a study performed in Veteran's Administration clinics demonstrated a strong patient preference for women-only centers (Bean–Mayberry et al., 2003); (Harpole et al., 2000; Phelan et al., 2000). This finding may support the possibility that populations with a high proportion of trauma-exposed women are more likely to prefer single-gender care settings and that the prevalence rates of PTSD could be higher in such settings. However, this has yet to be empirically proven.

The authors designed the current study to assess the prevalence of previously undetected screen positive, or suspected, PTSD and IPV in a dedicated women's headache treatment center population. This study replicates and

extends prior work (Peterlin et al., 2009) to a dedicated women's headache treatment setting. It also makes a novel contribution by estimating a 12-month rate of IPV among headache sufferers as prior studies only reported lifetime estimates. Researchers in the current study had several aims. First, they sought to estimate prevalence rates for PTSD and physical IPV (12-month and lifetime) in this all-female treatment setting. They expected that these rates would be higher than those reported previously. Secondly, they separately tested the associations between headache severity and PTSD and between headache severity and physical IPV. They theorized that the relationship between PTSD and headache severity would be stronger than that between physical IPV and headache severity because prior investigation supports that PTSD is more strongly linked to health problems than trauma per se (Schnurr & Green, 2004). In a study of primary care outpatients, trauma and extreme stress were linked to medical illness, but this effect was stronger for the subset of patients who had PTSD (Weisberg et al., 2002). PTSD itself appears to stress the affected individual, and traumatized women with PTSD have higher allostatic loads than traumatized women without PTSD (Glover, Stuber, & Poland, 2006).

MATERIALS AND METHODS

Study Sample

The authors conducted a cross-sectional survey study of a convenience sample of women presenting to a referral center for women with headache, based at an academic community hospital. Referred patients who presented with either migraine or tension variant headache to this treatment center in the greater Boston area between 2006 and 2007 were approached during regular operating hours of the clinic and asked if they would be willing to participate in a study on stress and headaches. The study was an unfunded exploratory study staffed by trained volunteer research assistants. Ninety-two female patients consented to complete the survey. Patients under the age of 18 years, pregnant women, and non-English speakers were eligible to participate, and were approached by a research assistant after an initial nursing intake form established eligibility. Participants gave written, signed informed consent administered by the research assistant prior to completing study questionnaires.

Data Collection

Those who chose to participate were asked basic demographic questions and also completed the Migraine Disability Assessment (MIDAS) questionnaire (Stewart, Lipton, Whyte, et al., 1999; Lipton et al., 2001) as part of their routine assessment. The MIDAS is a five-question self-report tool that

measures headache-related disability over the preceding 90 days in three domains: paid work/education, household/family matters, and social/leisure activities; the unit of measure is days (see Appendix). The MIDAS has been validated for use with both migraine and tension variant headaches (Stewart, Lipton, Kolodner, et al., 1999), and the highest possible score is 270 (Stewart et al., 2001). Respondents answered five questions about the impact of headache on their functioning, as measured in missed days of school/work, decrements in work and household productivity and interference in family, social, and leisure activities using a self-administered format. The total days of patient-reported disability were then summed into a point score. A score greater than 21 (days) indicated severe disability. The final MIDAS score has been validated as a measure of headache severity (Stewart, Lipton, Kolodner, et al., 1999; Lipton et al., 2001; Stewart et al., 2001).

PTSD was ascertained using a seven-item modified Breslau screening tool (Breslau et al., 1999) that has been validated in primary care outpatient settings against a clinician-administered structured interview measuring 17 core symptoms (CAPS; Kimerling et al., 2006) (Appendix). In the current study, the research assistant administered the questions to participants. The instrument was scored by summing all responses; the range of possible scores on this modified Breslau scale is 0–7. The cutoff score of 4 or greater for indicating PTSD was established using sensitivity, specificity, and likelihood ratios. A score over this cut point had a sensitivity (95% confidence interval) of 85% (73%, 97%) and a specificity of 84% (77%, 91%) (Kimerling et al., 2006).

Lifetime IPV was measured using the STaT ("slapped, threatened and throw"; Paranjape & Liebschutz, 2003), a three-question instrument developed by validation against a semi-structured interview, which had a sensitivity of 96% (90%, 100%); 89% (81%, 98%); and 64% (50%, 78%) for scores of 1, 2, and 3, respectively. The corresponding specificity is 75% (59%–91%), 100%, and 100%. The research assistant administered these questions. Endorsement of any item on the STaT was considered a positive screen for past or lifetime IPV. The semi-structured interview against which the STaT was validated included criteria measuring emotional control by an intimate partner. The research assistant also administered the Partner Violence Screen (PVS) to ascertain recent IPV. The PVS is a three-item instrument validated against the Index of Spousal Abuse (ISA) and the Conflict Tactics Scale (CTS) (Appendix). Compared with the ISA, the sensitivity of the PVS in detecting partner abuse was 64.5%; the specificity was 80.3% (Feldhaus et al., 1997). When compared with the CTS, sensitivity of the PVS was 71.4%; the specificity was 84.4%. Positive predictive values ranged from 51.3% to 63.4%, and negative predictive values ranged from 87.6% to 88.7%. For the PVS, any affirmative response was considered a positive screen when physical abuse was perpetrated by an intimate partner. Due to the need for brevity in this busy clinical setting, instruments expressly designed and validated to

measure emotional and sexual abuse were not included; researchers focused instead on measurement of physical abuse and the threat thereof.

Subsequently, the STaT responses were dichotomized into the past (lifetime) abuse variable, and the PVS responses formed the dichotomous variable for recent (12-month) IPV. Women who endorsed recent IPV received immediate safety planning and counseling, and those who reported any IPV or screened positive for PTSD were offered referrals to on-site mental health services.

A chart review was conducted to ascertain the MIDAS score (which was performed prior to the research interview by a clinician during the patient's medical visit on the same date of the study visit) along with the following variables: smoking, alcohol use history, marital status, and insurance type.

Statistical Analysis

Frequencies were calculated for survey and chart review data. To reduce potential bias from missing data on smoking and drinking habits, researchers performed multiple imputation modeling (Rubin, 1987). They first performed chi square tests on cross tabulations of categorical variables to test bivariate associations of the relationship between IPV and headache severity, and IPV and PTSD. In these analyses, headache severity (MIDAS score) was dichotomized into a categorical variable, using the cut point of 21 for these exploratory analyses to determine what percentage of women fell into the highest category of headache severity. Researchers calculated the needed sample size based on a theoretical 20-point clinical difference in MIDAS score between participants positive for, and those negative for PTSD. At the time the study began, no published literature was available on the likely magnitude of the relation of PTSD with headache. Using an *a priori* sample size calculator for multiple regression (Cohen et al., 2003), researchers estimated that with 80% power and $\alpha = 0.05$, a sample size of 70 would be sufficient to detect a theoretical 20-point difference in MIDAS score.

In all subsequent regression models, MIDAS was entered as a continuous variable to maximize statistical power. Next, researchers estimated unadjusted linear regression models to examine relationships between the dependent variable, MIDAS score, and PTSD, 12-month and lifetime IPV, age, insurance type, and tobacco and alcohol use. Statistical significance was set at $p = 0.05$. Covariates were retained in the final multivariable models based on their theoretical association with the dependent variable, headache severity as measured by the MIDAS (Peterlin, Nijjar, et al., 2011; Cripe et al., 2011; Peterlin et al., 2007; Coker et al., 2002; Domino & Haber, 1987; Aamodt et al., 2006; Schwartz et al., 1998; Diamond, 2007; Winter et al., 2012) as well as PTSD (Feldner, Babson, & Zvolensky, 2007; Kessler et al., 1995), and IPV (Gerber et al., 2005). Insurance type was included as a proxy for socioeconomic status (SES), which has also been associated with headache (Winter

et al., 2012), PTSD (Savoca & Rosenheck, 2000), and IPV (Cunradi, Caetano, & Schafer, 2002).

To determine whether recent and lifetime IPV could be entered into the same regression model, researchers estimated the correlation between the two. In the final step, they estimated three separate models: Model 1 examined the relationship between PTSD and headache severity (MIDAS), Model 2 examined recent IPV and headache, and Model 3 examined lifetime IPV and headache severity (MIDAS). Model fit was assessed by calculating an R^2 statistic. Analyses were conducted using SAS Version 9.2 [Cary, NC].

The study received Institutional Review Board approval at the Cambridge Health Alliance and VA Boston Healthcare System.

RESULTS

The mean age of the participants was 39 years (SD = 13.18, range 18–66 years), and the mean MIDAS score was 50.6 (SD = 52.46, range 0–240) (see Table 1). While participants came from a wide spectrum of socioeconomic strata, the majority of the participants in the study had private/commercial insurance (73%, 67/92). Most of the participants reported living with an intimate partner (73%, 67/92); 50% were married (46/92). Nine (9.8%) of the participants endorsed having experienced IPV within the preceding year, and 36.9% (34/92) reported lifetime IPV. The prevalence of a positive screen for PTSD was 28.3% (26/92).

PTSD was significantly associated with headache severity in bivariate analyses ($X^2 = 4.47$, $p = 0.03$). However, neither lifetime ($X^2 = 1.05$, $p = 0.3$) nor recent ($X^2 = 0.007$, $p = 0.9$) IPV was associated with headache severity in bivariate analyses. Researchers next examined the relationship

TABLE 1 Participant Characteristics, $n = 92$

Covariate	*N* (%) Mean (Standard Error)
PTSD screen positive (%) (cut point ≥4)	26 (28.3%)
MIDAS Score (mean) (SD)	50.6 (5.58)
Age in years (mean)	39.2 (1.37)
Had commercial health insurance (%)	68 (73.9%)
Married (%)	46 (50%)
Non-smoker (%)	78 (92%)*
Non-drinker (%)	47 (58%)**
Lifetime IPV, screen positive (%)	34 (36.9%)
Recent/12-month IPV screen positive (%)	9 (9.8%)

*Imputed data estimate = 430/460 (93.5%).
**Imputed data estimate = 295/460 (64.1%).

TABLE 2 Unadjusted and Multiple Linear Regression Results (dependent variable MIDAS score)

	Unadjusted			Adjusted		
Covariate	Coefficient	SE	*p*-Value	Coefficient	SE	*p*-Value
Model 1: Relation of PTSD to MIDAS						
PTSD	31.80	11.82	0.01	26.40	12.21	0.04
Age (years)	−0.004	0.42	0.99	0.015	0.42	0.72
Health insurance	−2.29	9.60	0.81	1.77	9.84	0.86
Tobacco use	18.78	12.74	0.15	27.03	14.07	0.06
Any alcohol use	−23.5	11.35	0.046	−22.16	11.97	0.07
Model 2: Relation of recent IPV to MIDAS						
Recent IPV	1.93	16.04	0.91	0.66	15.94	0.97
Age (years)	−0.004	0.42	0.99	0.09	0.43	0.84
Health insurance	−2.29	9.60	0.81	1.19	9.94	0.91
Tobacco use	18.78	12.74	0.15	26.71	13.74	0.06
Any alcohol use	−23.50	11.35	0.046	−28.02	11.79	0.02
Model 3: Relation of lifetime IPV to MIDAS						
Lifetime IPV	7.75	5.07	0.14	6.62	5.13	0.20
Age (years)	−0.004	0.42	0.99	0.06	0.42	0.90
Health insurance	−2.29	9.60	0.81	3.19	9.90	0.75
Tobacco use	18.78	12.74	0.15	25.36	13.59	0.07
Any alcohol use	−23.50	11.35	0.046	−26.19	11.69	0.03

between IPV and PTSD. Lifetime exposure to IPV was associated with PTSD ($X^2 = 21.02$, $p = 0.0001$), but recent IPV was not ($X^2 = 3.14$, $p = 0.21$).

In the multivariable regression models (Table 2), PTSD remained strongly associated with headache severity ($\beta = 34.12$, $p = 0.01$), after adjustment for potential confounding variables. This beta coefficient translated into an average of 34 more days of headache-related symptoms in a 90-day period and a statistically significant higher degree of headache severity for participants with PTSD than for those without PTSD. Lifetime and recent IPV were correlated ($p = 0.399$), and were therefore added to two separate regression models to examine their association with headache severity separately. Neither form of IPV was significantly associated with headache severity ($p = 0.05$) in these two separate regression models (Table 2). A gradient was observed for model fit in the R^2 results: Model 1 (PTSD and headache) $R^2 = 13.74\%$, Model 2 (lifetime IPV and headache) $R^2 = 11.04\%$, and Model 3 (recent IPV and headache) $R^2 = 9.2\%$.

DISCUSSION

The authors found that 28.3% of participants presenting to a women's headache treatment center screened positive for PTSD, which is in concert

with the range of 12.5% to 42.9% reported in other dedicated headache treatment settings (Peterlin et al., 2008, 2009), and also falls in the range of 11.8% to 33% reported in various primary care settings (Liebschutz et al., 2007; Stein et al., 2000; Alim et al., 2006). The study participants also exhibited high prevalence rates of both lifetime (36.9%) and recent IPV (9.8%) when compared to estimates of 24.3% and 2.7% of women in the general population (Black et al., 2011). The current results replicated the prevalence rates of lifetime IPV (27.3% to 34.8%) that have previously been reported (Cripe et al., 2011; Peterlin et al., 2007), and augment the existing literature by presenting a 12-month prevalence rate for these clinical settings. It is not surprising that the prevalence rates of IPV in this study exceeded population rates, as women with chronic pain report higher rates of exposure to abuse than those derived in population studies (Wuest et al., 2008; Nicolaidis & Liebschutz, 2009; Coker et al., 2002). The 12-month prevalence rate of IPV in this headache center population was comparable to 12-month rates of 8.6% reported among women in the same metropolitan area in primary care settings (McCloskey et al., 2005) where women frequently present with headache (Lipton et al., 2007).

These findings contradicted researchers original expectation that higher prevalence rates of suspected PTSD and IPV would be found in this all-female headache setting. One reason for this may have been the high proportion of women represented in the previously published estimates in treatment settings for both male and female patients. It is also conceivable that patients presenting to an all-female clinical setting do not differ on trauma exposure and potential susceptibility to PTSD. However, researchers may also have underestimated the true prevalence of IPV in the current population by measuring primarily one domain of IPV (physical, not sexual or psychological). Although neither of the two prior studies of IPV and headache cited previously included emotional abuse estimates, both measured physical and sexual abuse (Cripe et al., 2011; Peterlin et al., 2007). It remains possible that had sexual abuse rates been measured in the current study, the observed prevalence rates of IPV may have been higher. It is also plausible that headache is more strongly associated with psychological abuse, and that the authors underestimated the magnitude of this relationship by using limited measures.

The current study also reaffirmed the findings from prior work demonstrating increased headache impact scores in participants with PTSD using a different measure of headache severity, the HIT-6 (Peterlin et al., 2009). In the current study, the mean MIDAS headache severity score of 50, well above the cut point of 21, suggests that study participants were experiencing a high degree of headache-related disability. PTSD status was statistically significantly associated with MIDAS score; participants who screened positive for PTSD had headache severity scores that were 34 points higher on average than those negative for PTSD, controlling for covariates.

This translates into 34 more days of headache-related disruption in activities over a 90-day period, or disabling headache more than 30% of the time. This degree of headache severity could affect treatment considerations. Within this population, IPV was not significantly associated with headache severity at the $p < 0.05$ level. Yet, patients who reported exposure to recent and lifetime IPV scored, respectively, 1.6 and 9 points higher on the MIDAS, suggesting that, despite lack of statistical significance, these women exhibited greater headache severity and more days of adverse impact from headache. This trend appears particularly salient for women who have experienced lifetime IPV.

These results may reflect that the development of PTSD (or suspected PTSD) is a more important factor in determining headache severity than the trauma alone (in this study, IPV exposure), as suggested previously (Schnurr & Green, 2004; Kimerling, Clum, & Wolfe, 2000). The normative response to trauma (including IPV) is resilience. Most trauma survivors recover naturally. In the minority of cases where PTSD develops, it puts "load" on the system and increases risk for headache over and above IPV exposure. Headache patients with PTSD may have higher allostatic load with accompanying physiologic changes that manifest as greater headache severity. These findings are consonant with the growing body of research documenting PTSD as a primary mechanism contributing to the relationship between IPV and physical health problems (Woods et al., 2008).

This referral population was relatively well insured, underscoring the fact that IPV affects women from all socioeconomic strata. This finding emphasizes that clinicians routinely caring for women with headaches must be attuned to the potential presence of IPV during their patient encounters. Clinicians and care teams should be prepared to offer resources and referrals when patients do disclose abuse. Additionally, because both headache and PTSD are conditions that occur more frequently in women, healthcare providers who routinely treat women need to be cognizant of the association between headache and PTSD, especially in women with chronic headaches who tend to have higher headache severity scores in the range reported by this study.

Limitations

In the current data, headache was not classified by formal International Classification of Headache Disorders (ICHD-2) diagnostic guidelines, as have been used in other studies (Peterlin et al., 2007, 2008, 2009; Peterlin, Rosso, et al., 2011; Cripe et al., 2011), because the authors felt that this information was often not available to primary care providers seeing women patients, and because the MIDAS has been validated in both tension variant and migraine headaches (Stewart, Lipton, Kolodner, et al., 1999). However, this limited the ability to compare current findings to those of other studies that have

stratified their results by headache type. In addition, data on head injury and depression were not available, both of which could be confounders due to their potential associations with physical trauma as well as headache. Further, statistical power to detect as statistically significant meaningful differences in headache severity by IPV exposure may have been limited by the small sample size. The small numbers of women who reported recent IPV in particular may have limited the researchers' ability to detect meaningful differences in headache severity. Additionally, as stated above, the PTSD screen did not determine whether the symptoms queried were anchored to IPV exposure specifically. Thus, it is possible that the positive PTSD screening data ascertained with this measurement was secondary to other types of trauma (including child abuse and sexual assault) not measured in this study. Other limitations included the cross-sectional design of the study, which prohibited assessment of the timing of the IPV and PTSD in relation to the development of headaches, and potential challenges to generalizeability, given that this was a clinical convenience sample from a unique treatment setting to which women with very severe headaches were referred. Participation bias was also possible; researchers did not collect data on all women approached in the clinic and on women who did not participate in the study, to be able to assess potential selection and participation biases. Finally, the administration of potentially sensitive trauma-related measures by a research assistant may have increased the possibility of reporting bias in the sample.

CONCLUSION

This study, conducted in a women's headache specialty center, contributes additional evidence to the existing literature linking PTSD to adverse health outcomes, notably headache severity. While the study did not confirm the prior expectation that prevalence rates of PTSD and IPV could be higher in a dedicated women's headache population, the restricted definition of IPV may have underestimated the true prevalence of all forms of IPV in this population. This study does contribute prevalence estimates using well-validated IPV measures, thus confirming prior observations that IPV is a salient issue for women with headaches. The data on rates of recent IPV contribute new evidence to the imperative that providers treating women with headaches be aware of recent physical abuse in their patient populations. Clinicians practicing in dedicated headache treatment settings may have a unique opportunity to address abuse history (Schulman, 2010).

Future work examining interventions for abused women with headache would be an important contribution to improving care for this patient population that is often treated not only by neurologists but primary care clinicians working in women's health. The strong association between headache severity and PTSD presents a potential multi-disciplinary treatment path where traditional headache treatment is augmented with empirically supported

treatments for PTSD, such as cognitive behavioral therapy. Informed by understanding that PTSD and chronic pain are linked, researchers have demonstrated that integrated treatment is beneficial for participants with these co-morbidities (Otis et al., 2009; Shipherd et al., 2003). Even effective treatment of PTSD alone may be helpful in improving chronic pain (Otis et al., 2003), suggesting that future intervention studies in populations with severe and chronic headaches could yield similar benefit.

REFERENCES

Aamodt, A. H., L. J. Stovner, K. Hagen, G. Brathen, and J. Zwart. 2006. Headache prevalence related to smoking and alcohol use. The Head-HUNT Study. *Eur J Neurol* 13(11):1233–8.

Alim, T. N., E. Graves, T. A. Mellman, N. Aigbogun, E. Gray, W. Lawson, and D. S. Charney. 2006. Trauma exposure, posttraumatic stress disorder and depression in an African-American primary care population. *J Natl Med Assoc* 98(10):1630–6.

American Hospital Association. 2000. AHA hospital statistics 00. Chicago: Health Forum LLC.

American Psychological Association. 2000. *Diagnostic and statistical manual of mental disorders*, revised 4th ed. (DSM-IV). Washington, DC: Author.

Anda, R., G. Tietjen, E. Schulman, V. Felitti, and J. Croft. 2010. Adverse childhood experiences and frequent headaches in adults. *Headache* 50(9):1473–81.

Bean–Mayberry, B. A., C. C. Chang, M. A. McNeil, J. Whittle, P. M. Hayes, and S. H. Scholle. 2003. Patient satisfaction in women's clinics versus traditional primary care clinics in the Veterans Administration. *J Gen Intern Med* 18(3):175–81.

Black, M. C., K. C. Basile, M. J. Breiding, S. G. Smith, M. L. Walters, M. T. Merrick, et al. 2011. National Intimate Partner and Sexual Violence Survey (NISVS): 2010 summary report. Retrieved March 14, 2012, from http://www.cdc.gov/ViolencePrevention/pdf/NISVS_Report2010-a.pdf

Breslau, N., E. L. Peterson, R. C. Kessler, and L. R. Schultz. 1999. Short screening scale for DSM-IV posttraumatic stress disorder. *Am J Psychiat* 156(6):908–11.

Campbell, J., A. S. Jones, J. Dienemann, J. Kub, J. Schollenberger, P. O'Campo, et al. 2002. Intimate partner violence and physical health consequences. *Arch Intern Med* 162(10):1157–63.

Cohen, J., P. Cohen, S. G. West, and L. S. Aiken. 2003. *Applied multiple regression/correlation analysis for the behavioral sciences*. 3rd ed. Mahwah, NJ: Lawrence Earlbaum Associates.

Coker, A. L., K. E. Davis, I. Arias, S. Desai, M. Sanderson, H. M. Brandt, and P. H. Smith. 2002. Physical and mental health effects of intimate partner violence for men and women. *Am J Prev Med* 23(4):260–8.

Cripe, S. M., S. E. Sanchez, B. Gelaye, E. Sanchez, and M. A. Williams. 2011. Association between intimate partner violence, migraine and probable migraine. *Headache* 51(2):208–19.

Cunradi, C. B., R. Caetano, and J. Schafer. 2002. Socioeconomic predictors of intimate partner violence among white, black, and Hispanic couples in the United States. *J Fam Violence* 17(4):377–389.

De Leeuw, R., J. E. Schmidt, and C. R. Carlson. 2005. Traumatic stressors and posttraumatic stress disorder symptoms in headache patients. *Headache* 45(10):1365–74.

Dennis, M. F., A. M. Flood, V. Reynolds, G. Araujo, C. P. Clancy, J. C. Barefoot, and J. C. Beckham. 2009. Evaluation of lifetime trauma exposure and physical health in women with posttraumatic stress disorder or major depressive disorder. *Violence Against Women* 15(5):618–27.

Diamond, M. 2007. The impact of migraine on the health and well-being of women. *J Women's Health* 16(9):1269–80.

Domino, J. V., and J. D. Haber. 1987. Prior physical and sexual abuse in women with chronic headache: Clinical correlates. *Headache* 27(6):310–4.

Dutton, M. A., B. L. Green, S. I. Kaltman, D. M. Roesch, T. A. Zeffiro, and E. D. Krause. 2006. Intimate partner violence, PTSD, and adverse health outcomes. *J Interpers Violence* 21(7):955–68.

Elwood, L. S., N. L. Williams, B. O. Olatunji, and J. M. Lohr. 2007. Interpretation biases in victims and non-victims of interpersonal trauma and their relation to symptom development. *J Anxiety Disord* 21(4):554–67.

Feldhaus, K. M., J. Koziol–McLain, H. L. Amsbury, I. M. Norton, S. R. Lowenstein, and J. T. Abbott. 1997. Accuracy of 3 brief screening questions for detecting partner violence in the emergency department. *JAMA* 277(17):1357–61.

Feldner, M. T., K. A. Babson, and M. J. Zvolensky. 2007. Smoking, traumatic event exposure, and Posttraumatic stress: A critical review of the empirical literature. *Clin Psychol Rev* 27(1):14–45.

Gerber, M. R., M. L. Ganz, E. Lichter, C. M. Williams, and L. A. McCloskey. 2005. Adverse health behaviors and the detection of partner violence by clinicians. *Arch Intern Med* 165(9):1016–21.

Gill, J. M., S. L. Szanton, and G. G. Page. 2005. Biological underpinnings of health alterations in women with PTSD: A sex disparity. *Biol Res Nurs* 7(1):44–54.

Glover, D. A., M. Stuber, and R. E. Poland. 2006. Allostatic load in women with and without PTSD symptoms. *Psychiatry* 69(3):191–203.

Golding, J. M. 1999. Sexual assault history and headache: Five general population studies. *J Nerv Ment Dis* 187(10):624–9.

Harpole, L. H., E. A. Mort, K. M. Freund, J. Orav, and T. A. Brennan. 2000. A comparison of the preventive health care provided by women's health centers and general internal medicine practices. *J Gen Intern Med* 15(1):1–7.

Heron, S. L., M. P. Thompson, E. Jackson, and N. J. Kaslow. 2003. Do responses to an intimate partner violence screen predict scores on a comprehensive measure of intimate partner violence in low-income black women? *Ann Emerg Med* 42(4):483–91.

Kessler, R. C., A. Sonnega, E. Bromet, M. Hughes, and C. B. Nelson. 1995. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiat* 52(12):1048–60.

Kimerling, R., G. A. Clum, and J. Wolfe. 2000. Relationships among trauma exposure, chronic posttraumatic stress disorder symptoms, and self-reported health in women: Replication and extension. *J Trauma Stress* 13(1):115–28.

Kimerling, R., P. Ouimette, A. Prins, P. Nisco, C. Lawler, R. Cronkite, and R. H. Moos. 2006. Brief report: Utility of a short screening scale for DSM-IV PTSD in primary care. *J Gen Intern Med* 21(1):65–7.

Liebschutz, J., R. Saitz, V. Brower, T. M. Keane, C. Lloyd–Travaglini, T. Averbuch, and J. H. Samet. 2007. PTSD in urban primary care: High prevalence and low physician recognition. *J Gen Intern Med* 22(6):719–26.

Lipton, R. B., M. E. Bigal, M. Diamond, F. Freitag, M. L. Reed, and W. F. Stewart. 2007. Migraine prevalence, disease burden, and the need for preventive therapy. *Neurology* 68(5):343–9.

Lipton, R. B., W. F. Stewart, J. Sawyer, and J. G. Edmeads. 2001. Clinical utility of an instrument assessing migraine disability: The Migraine Disability Assessment (MIDAS) questionnaire. *Headache* 41(9):854–61.

Lowe, B., K. Kroenke, R. L. Spitzer, J. B. Williams, M. Mussell, M. Rose, et al. 2011. Trauma exposure and posttraumatic stress disorder in primary care patients: Cross-sectional criterion standard study. *J Clin Psychiat* 72(3): 304–12.

McCloskey, L. A., E. Lichter, M. L. Ganz, C. M. Williams, M. R. Gerber, R. Sege, et al. 2005. Intimate partner violence and patient screening across medical specialties. *Acad Emerg Med* 12(8):712–22.

Nicolaidis, C., and J. Liebschutz. 2009. Chronic physical symptoms in survivors of intimate partner violence. In *Intimate partner violence: A health-based perspective*, ed. C. Mitchell and D. Anglin, 133–45. New York: Oxford University Press.

Otis, J. D., T. M. Keane, and R. D. Kerns. 2003. An examination of the relationship between chronic pain and Posttraumatic stress disorder. *J Rehabil Res Dev* 40(5):397–405.

Otis, J. D., T. M. Keane, R. D. Kerns, C. Monson, and E. Scioli. 2009. The development of an integrated treatment for veterans with comorbid chronic pain and posttraumatic stress disorder. *Pain Med* 10(7):1300–11.

Paranjape, A., and J. Liebschutz. 2003. STaT: A three-question screen for intimate partner violence. *J Womens Health (Larchmt)* 12(3):233–9.

Paras, M. L., M. H. Murad, L. P. Chen, E. N. Goranson, A. L. Sattler, K. M. Colbenson, et al. 2009. Sexual abuse and lifetime diagnosis of somatic disorders: A systematic review and meta-analysis. *JAMA* 302(5):550–61.

Peterlin, B. L., S. S. Nijjar, and G. E. Tietjen. 2011. Posttraumatic stress disorder and migraine: Epidemiology, sex differences, and potential mechanisms. *Headache* 51(6):860–8.

Peterlin, B. L., A. L. Rosso, F. D. Sheftell, D. J. Libon, J. M. Mossey, and K. R. Merikangas. 2011. Posttraumatic stress disorder, drug abuse and migraine: New findings from the National Comorbidity Survey Replication (NCS-R). *Cephalalgia* 31(2):235–44.

Peterlin, B. L., G. E. Tietjen, J. L. Brandes, S. M. Rubin, E. Drexler, J. R. Lidicker, and S. Meng. 2009. Posttraumatic stress disorder in migraine. *Headache* 49(4): 541–51.

Peterlin, B. L., G. Tietjen, S. Meng, J. Lidicker, and M. Bigal. 2008. Posttraumatic stress disorder in episodic and chronic migraine. *Headache* 48(4):517–22.

Peterlin, B. L., T. Ward, J. Lidicker, and M. Levin. 2007. A retrospective, comparative study on the frequency of abuse in migraine and chronic daily headache. *Headache* 47(3):397–401.

Phelan, E. A., W. Burke, R. A. Deyo, T. D. Koepsell, and A. Z. LaCroix. 2000. Delivery of primary care to women. Do women's health centers do it better? *J Gen Intern Med* 15(1):8–15.

Raphael, K. G., and C. S. Widom. 2010. Posttraumatic stress disorder moderates the relation between documented childhood victimization and pain 30 years later. *Pain* 152(1): 163–9.

Rasmusson, A. M., P. P. Schnurr, Z. Zukowska, E. Scioli, and D. E. Forman. 2010. Adaptation to extreme stress: Posttraumatic stress disorder, neuropeptide Y and metabolic syndrome. *Exp Biol Med (Maywood)* 235(10):1150–62.

Romans, S., C. Belaise, J. Martin, E. Morris, and A. Raffi. 2002. Childhood abuse and later medical disorders in women. An epidemiological study. *Psychother Psychosom* 71(3):141–50.

Rubin, D. B. 1987. *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc.

Savoca, E., and R. Rosenheck. 2000. The civilian labor market experiences of Vietnam-era veterans: The influence of psychiatric disorders. *J Ment Health Policy Econ* 3(4):199–207.

Schnurr, P.P., and B.L. Green. 2004. *Trauma and health: Physical health consequences of exposure to extreme stress*. Washington, DC: American Psychological Association.

Schnurr, P. P., and M. K. Jankowski. 1999. Physical health and posttraumatic stress disorder: Review and synthesis. *Semin Clin Neuropsychiatry* 4(4):295–304.

Schulman, E. 2010. *Abuse screening should be part of the neurology exam*. Retrieved March 24, 2012, from http://www.aan.com/advocacy/advocates/files/attachments/Schulman%20-%20World%20Neurology.pdf

Schwartz, B. S., W. F. Stewart, D. Simon, and R. B. Lipton. 1998. Epidemiology of tension-type headache. *JAMA* 279(5):381–3.

Sharp, T. J., and A. G. Harvey. 2001. Chronic pain and posttraumatic stress disorder: Mutual maintenance? *Clin Psychol Rev* 21(6):857–77.

Shipherd, J. C., J. G. Beck, J. L. Hamblen, J. M. Lackner, and J. B. Freeman. 2003. A preliminary examination of treatment for posttraumatic stress disorder in chronic pain patients: A case study. *J Trauma Stress* 16(5):451–7.

Stein, M. B., J. R. McQuaid, P. Pedrelli, R. Lenox, and M. E. McCahill. 2000. Posttraumatic stress disorder in the primary care medical setting. *Gen Hosp Psychiat* 22(4):261–9.

Stewart, W. F., R. B. Lipton, A. J. Dowson, and J. Sawyer. 2001. Development and testing of the Migraine Disability Assessment (MIDAS) questionnaire to assess headache-related disability. *Neurology* 56(6 suppl 1):S20–8.

Stewart, W. F., R. B. Lipton, K. Kolodner, J. Liberman, and J. Sawyer. 1999. Reliability of the migraine disability assessment score in a population-based sample of headache sufferers. *Cephalalgia* 19(2):107–14; discussion 74.

Stewart, W. F., R. B. Lipton, J. Whyte, A. Dowson, K. Kolodner, J. N. Liberman, and J. Sawyer. 1999. An international study to assess reliability of the Migraine Disability Assessment (MIDAS) score. *Neurology* 53(5):988–94.

Stokes, M., W. J. Becker, R. B. Lipton, S. D. Sullivan, T. K. Wilcox, L. Wells, et al. 2011. Cost of health care among patients with chronic and episodic migraine in Canada and the USA: Results from the International Burden of Migraine Study (IBMS). *Headache* 51(7):1058–77.

Tam, S., D. Joyce, M. R. Gerber, and A. Tan. 2010. Head and neck injuries in adult victims of intimate-partner violence. *J Otolaryngol Head Neck Surg* 39(6):737–43.

Tietjen, G. E., J. L. Brandes, B. L. Peterlin, A. Eloff, R. M. Dafer, M. R. Stein, et al. 2009a. Childhood maltreatment and migraine (part I). Prevalence and adult revictimization: A multicenter headache clinic survey. *Headache* 50(1): 20–31.

Tietjen, G. E., J. L. Brandes, B. L. Peterlin, A. Eloff, R. M. Dafer, M. R. Stein, et al. 2009b. Childhood maltreatment and migraine (part II). Emotional abuse as a risk factor for headache chronification. *Headache* 50(1): 32–41.

Tolin, D. F., and E. B. Foa. 2006. Sex differences in trauma and posttraumatic stress disorder: A quantitative review of 25 years of research. *Psychol Bull* 132(6):959–92.

Walling, M. K., R. C. Reiter, M. W. O'sHara, A. K. Milburn, G. Lilly, and S. D. Vincent. 1994. Abuse history and chronic pain in women: I. Prevalences of sexual abuse and physical abuse. *Obstet Gynecol* 84(2):193–9.

Weisberg, R. B., S. E. Bruce, J. T. Machan, R. C. Kessler, L. Culpepper, and M. B. Keller. 2002. Nonpsychiatric illness among primary care patients with trauma histories and posttraumatic stress disorder. *Psychiatr Serv* 53(7):848–54.

Winter, A. C., K. Berger, J. E. Buring, and T. Kurth. 2012. Associations of socioeconomic status with migraine and non-migraine headache. *Cephalalgia* 32(2): 159–170.

Woods, S. J., R. J. Hall, J. C. Campbell, and D. M. Angott. 2008. Physical health and posttraumatic stress disorder symptoms in women experiencing intimate partner violence. *J Midwifery Wom Heal* 53(6):538–46.

Wuest, J., M. Ford–Gilboe, M. Merritt–Gray, C. Varcoe, B. Lent, P. Wilk, and J. Campbell. 2009. Abuse-related injury and symptoms of posttraumatic stress disorder as mechanisms of chronic pain in survivors of intimate partner violence. *Pain Medicine* 10(4):739–47.

Wuest, J., M. Merritt–Gray, M. Ford–Gilboe, B. Lent, C. Varcoe, and J. C. Campbell. 2008. Chronic pain in women survivors of intimate partner violence. *J Pain* 9(11):1049–57.

APPENDIX

MIDAS (Migraine Disability Assessment) Test (Stewart, et al. 1999)

(1) On how many days in the last 3 months did you miss work or school because of your headaches?
(2) How many days in the last 3 months was your productivity at work or school reduced by half or more because of your headaches? (Do not include days you counted in question 1 where you missed work or school.)
(3) On how many days in the last 3 months did you not do household work (such as housework, home repairs and maintenance, shopping, caring for children and relatives) because of your headaches?
(4) How many days in the last 3 months was your productivity in household work reduced by half of more because of your headaches? (Do not include days you counted in question 3 where you did not do household work.)

(5) On how many days in the last 3 months did you miss family, social or leisure activities because of your headaches?

Total (Questions 1–5)

STaT (Paranjape & Liebschutz, 2003)

(1) Have you ever been in a relationship where your partner has pushed or slapped you?
(2) Have you ever been in a relationship where your partner has thrown, broken or punched things?
(3) Have you ever been in a relationship where your partner has threatened you with violence?

Partner Violence Screen (PVS; Feldhaus et al., 1997)

(1) Have you been hit, kicked, punched or otherwise hurt by someone in the past year? If so, by whom?
(2) Do you feel safe in your current relationship?
(3) Is there a partner from a previous relationship who is making you feel unsafe now?

Short Screening Scale for PTSD (Kimerling et al., 2006)

In your life, have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month . . .

(1) Did you avoid being reminded of this experience by staying away from certain places, people, or activities?
(2) Did you lose interest in activities that were once important or enjoyable?
(3) Did you begin to feel more isolated or distant from other people?
(4) Did you find it hard to have love or affection for other people?
(5) Did you begin to feel that there was no point in planning for the future?
(6) After this experience were you having more trouble than usual falling asleep or staying asleep?
(7) Did you become jumpy or get easily startled by ordinary noises or movements?

Responses are YES = 1 or NO = 0. The scale is scored by summing all responses. Scale scores may range from 0 to 7.

Published in final edited form as:
J Emerg Med. 2012 June ; 42(6): 719–726. doi:10.1016/j.jemermed.2010.11.038.

A BUDGET IMPACT ANALYSIS OF RAPID HUMAN IMMUNODEFICIENCY VIRUS SCREENING IN VETERANS ADMINISTRATION EMERGENCY DEPARTMENTS

Risha Gidwani, DRPH[*,†], **Matthew Bidwell Goetz, MD**[*,‡], **Gerald Kominski, PHD**[†], **Steven Asch, MD, MPH**[*,‡], **Kristin Mattocks, PHD**[§,‖], **Jeffrey H. Samet, MD, MA, MPH**[¶], **Amy Justice, MD, PHD**[§,‖], **Neel Gandhi, MD**[**], and **Jack Needleman, PHD**[†]

[*] QUERI-HIV/HEPATITIS C, VA Greater Los Angeles Healthcare System, Los Angeles, California

[†] UCLA School of Public Health, Los Angeles, California

[‡] David Geffen School of Medicine at UCLA, Los Angeles, California

[§] West Haven VA Healthcare System, West Haven, Connecticut

[‖] Yale University School of Medicine, New Haven, Connecticut

[¶] Boston University School of Medicine, Boston, Massachusetts

[**] Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, New York

Abstract

Background—Human immunodeficiency virus (HIV) screening is cost-effective and recommended in populations with low disease prevalence. However, because screening is not cost-saving, its financial feasibility must be understood.

Study Objectives—We forecast the costs of two Emergency Department-based HIV testing programs in the Veterans Administration: 1) implementing a non-targeted screening program and providing treatment for all patients thusly identified (Rapid Testing); and 2) treating patients identified due to late-stage symptoms (Usual Care); to determine which program was the most financially feasible.

Methods—Using a dynamic decision-analysis model, we estimated the financial impact of each program over a 7-year period. Costs were driven by patient disease-severity at diagnosis, measured by CD4+ category, and the proportion of patients in each disease-severity category. Cost per CD4+ category was modeled from chart review and database analysis of treatment-naïve HIV-positive patients. Distributions of CD4+ counts differed in patients across the Rapid Testing and Usual Care arms.

Results—A non-targeted Rapid Testing program was not significantly more costly than Usual Care. Although Rapid Testing had substantial screening costs, they were offset by lower inpatient expenses associated with earlier identification of disease. Assuming an HIV prevalence of 1% and 80% test acceptance, the cost of Rapid Testing was $1,418,088, vs. $1,320,338 for Usual Care (p = 0.5854). Results support implementation of non-targeted rapid HIV screening in integrated systems. Conclusions: This analysis adds a new component of support for HIV screening by demonstrating that rapid, non-targeted testing does not cost significantly more than a diagnostic testing approach.

Reprint Address: Risha Gidwani, DRPH, rishagidwani@yahoo.com.

Keywords

HIV; screening; budget impact; Rapid Testing; Emergency Department

INTRODUCTION

It is estimated that there are 1–1.2 million people living in the United States with human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS), with 21% of them unaware of their disease status (1,2). Many HIV-positive persons are identified only when they develop symptoms, indicating that they are severely immunosuppressed and less likely to respond optimally to antiretroviral therapy. Screening programs that diagnose patients early and offer them treatment can therefore substantially improve health outcomes (3).

Recently, four separate analyses have indicated that HIV screening is cost-effective from a societal perspective compared to no screening and compared to current practice, prompting the Centers for Disease Control and Prevention (CDC) and the American College of Physicians to encourage routine HIV screening as a part of normal medical practice (4–9). However, although HIV screening has been found to be cost-effective, it is not cost-saving. Therefore, even though HIV screening is economically justifiable, it may not be a financially viable option for an organization. A Budget Impact Analysis (BIA) can provide information about the financial feasibility of a program by examining the economic value of the health investment and the resources needed for its implementation (10).

This study examines the budget impact of implementing a new HIV testing program in a Veterans Health Administration (VA) Emergency Department (ED) vs. the financial impact of following standard care. The standard care program examined is Usual Care, which in the ED setting involves blood-based diagnostic testing, or testing a patient when (s)he presents with symptoms suggestive of HIV/AIDS. The new program examined is Rapid Testing, or offering an HIV test to any previously untested patient in the ED, regardless of risk factors or symptoms (non-targeted testing). The rapid test in question is the OraQuick Advance HIV-1/2 Oral Specimen Collection Device (OraSure Technologies, Bethlehem, PA), a Food and Drug Administration-approved point-of-care test. Of particular interest for ED settings, the results of HIV rapid tests are available within 1 h, compared to the 24–48 h turnaround for blood-based testing.

This analysis estimates the number of people who would be identified as disease-positive through Rapid Testing at various levels of program intensity and HIV prevalence rates. The model forecasts treatment costs for patients in the Rapid Test program and compares them to the expenses incurred by these same patients were they to be identified at later stages of disease through Usual Care. The program with the lowest overall costs, both in terms of implementation costs and cost offsets, represents the most economically-efficient strategy.

MATERIALS AND METHODS

We built a first-order stochastic decision-analysis model to determine the costs of treating a hypothetical cohort of patients identified due to a Rapid Testing screening program vs. costs of treating this same cohort of patients were they identified due to symptoms of disease (Usual Care). Cohort size indicates the number of HIV-positive persons identified by the screening programs in the ED. The ED was chosen as the screening site due to its ability to access an otherwise difficult-to-reach population and its predicted higher prevalence of HIV infection compared to primary care locations (11,12). This analysis used cost and economic data from an urban VA ED and Infectious Disease clinic located in a major metropolitan

area. The VA was chosen due to the extensive data available from its cost databases and electronic medical records. Institutional Review Board approval was obtained from the VA system.

Due to the fact that the VA generally keeps patients for life, this analysis assumed all patients would be identified at some point in the system due to their symptoms. Therefore, sizes for the Rapid Test and Usual Care cohorts were identical. We modeled a variety of cohort sizes to reflect uncertainty in the number of patients offered testing, the percentage of patients accepting testing, and the prevalence of HIV (Table 1). This analysis evaluated the financial impact of three different offer rates: five per business day, 10 per business day, or universal. The former two rates were used to reflect realistic levels of testing that can occur in environments with existing capacity constraints; the latter was used in accordance with CDC recommendations. The model assumes HIV prevalence is independent of test acceptance.

Costs of Rapid Testing include both program implementation and treatment of disease. Costs of Usual Care include disease-treatment costs only, but health care utilization is more intensive in this population due to late identification of disease. Treatment costs were calculated for the following budgets: inpatient, outpatient, pharmacy, and global (where global indicates the sum of the former three sub-budgets). All treatment costs were assumed to be dependent upon disease severity at diagnosis (measured by CD4+ count). Disease severity was noted by the following four categories: CD4+ < 50, CD4+ 50–199, CD4+ 200–350, and CD4+ > 350 cells/mm^3 at diagnosis (where low CD4+ values indicate higher disease severity). Costs were projected over a time horizon of 7 years. A horizon of 7 years is longer than usual for BIAs; however, we felt this longer horizon was necessary to capture the delay between when patients are diagnosed through screening and when they begin accruing costs of treatment.

Treatment cost was a function of cost per disease-severity category and the proportion of cohort members in each disease-severity category. Estimated costs were calculated on an annual basis for each budget using 2007 dollars. The following equation was used to model the total estimated costs of treating disease:

$$Cj = \sum_{i=1}^{4} \sum_{t=0}^{6} (C_{jit} * N_{jit})$$

Where:

C_j is the average cost per case in sub-budget category j

i is the disease-severity category (CD4+) assigned at time of diagnosis

t is the year in which costs are accrued

C is the average cost per case of disease

N is the number of people

Cost per disease-severity category was determined by examining utilization patterns of HIV-positive patients and allocating VA-specific direct medical cost values to this utilization. VA-specific costs were used, following best practice recommendations for conducting BIAs. (10). Utilization data came from chart review of all treatment-naïve patients diagnosed and treated at this site from Fiscal Year (FY) 2000 to FY 2007 (n = 112). Although this number is small, the analysis required the use of these actual data, due to the dearth of literature available on treatment-naïve patients and the fact that data from treatment-experienced

patients would substantially overestimate treatment costs. Modeling HIV-related utilization over time also poses significant difficulties, given that drug treatment efficacy decreases over time, is influenced by patient adherence, and that the relationship between adherence and consequences of much highly active anti-retroviral therapy (HAART) is not fully understood (13). For example, partial adherence increases the risk of drug resistance more so than complete non-adherence (13).

Only HIV-related utilization was included in the model. All inpatient stays with HIV as a primary or secondary diagnosis were reviewed by two infectious disease physicians who had treated the majority of these patients; only those inpatient stays unanimously deemed to be due to HIV were included in this analysis.

Once the cost per disease-severity category was calculated, we populated the model with the program-specific CD4+ distributions. As Usual Care patients are diagnosed at later, more advanced stages of disease than are patients identified through Rapid Testing, the disease-severity distributions differed across the programs; a larger proportion of Usual Case patients had lower CD4+ counts at diagnosis. The CD4+ distribution for the Usual Care program was determined using patients from the national Veterans Aging Cohort Study (VACS) Virtual Cohort, which contains clinical data on all HIV-infected veterans in the country (14). Disease severity at diagnosis was obtained for VACS Virtual Cohort members who had ED visits in a VA facility before their HIV diagnosis (n = 3355), which represents the Usual Care patients in this model.

We used a back-calculation scheme to determine the CD4+ distribution for the Rapid Testing program. Annual CD4+ decline in the absence of treatment can be estimated by the following equation: CD4+^½ =− .0584 – 0.918 [log(vRNA/1000)] (15). Entering patients' viral load at diagnosis, CD4+ count at diagnosis, as well as the length of time elapsed between first ED visit and HIV diagnosis into the equation yields an estimate of patients' CD4+ count had they been diagnosed with HIV at the time of first ED visit (a Rapid Testing approach). For example, a patient presenting 2 years after his first ED visit with an initial CD4 count of 150 cells/mm^3 and viral load of 100,000 copies would be back-calculated to have had a CD4 of 292 cells/mm^3 had he been diagnosed at the time of initial ED visit (Table 2).

As identified HIV-positive patients, persons in the Rapid Testing arm begin accruing treatment costs immediately. However, it takes some time for Usual Care patients to display symptoms of disease, be diagnosed, and therefore start incurring treatment costs. Data from the VACS Virtual Cohort indicate a mean 1.25-year delay between the first ED visit and diagnosis of HIV. All estimated costs in the Usual Care program were adjusted back 1.25 years to reflect this delay to diagnosis. Over the entire analysis, the Rapid Test program therefore had 7 years of data, whereas the Usual Care program had 5.75 years of data.

In the outpatient and pharmacy models, cost values associated with the Usual Care and Rapid Test programs were identical. The model assumed drug costs and office visits were not driven by method of HIV diagnosis, but simply by the severity of patient illness. However, cost values for the inpatient and thus global models differed across the Usual Care and Rapid Testing programs; inpatient utilization, and cost, was dependent upon whether a patient was diagnosed through Rapid Testing or Usual Care. Inpatient stays within 3 months of HIV diagnosis were excluded from the Rapid Testing costs. These hospitalizations represent patients who were diagnosed due to opportunistic infections; their exclusion was more indicative of the clinical scenario of Rapid Testing patients, who are not diagnosed due to opportunistic infections and therefore would not have experienced such inpatient stays.

Once the model was populated with the cost per CD4+ category and the distributions of CD4+ counts, we ran the analysis and employed statistical techniques to mitigate the noise associated with small samples. Probabilistic sensitivity analyses (with 10,000 Monte Carlo iterations) were run to gain a more accurate estimate of cost per CD4+ category, with each iteration randomly sampling from each distribution of cost per CD4+ category. This re-sampling increases the likelihood that outlier cost inputs would not have an undue influence on cost estimates. Model analyses were conducted using TreeAge Pro 2008 software (TreeAge Software Inc., Williamstown MA).

Usual Care program costs were directly obtained from model results. Rapid Testing program costs were obtained by adding program implementation costs to model results. Implementation costs were determined by micro-cost estimates of resources used in offering and conducting HIV tests, including testing kits and labor. Time-and-motion data were obtained by direct observation of HIV rapid testing in the ED. An applied rate of 85% was used for all salary calculations. Two-tailed *t*-tests of significance were run on the total costs of each program to determine if the overall 7-year global costs of the Rapid Test program were significantly different from the Usual Care program. Tests of significance were run using STATA version 9.2 (StataCorp LP, College Station, TX).

To summarize, the following assumptions were made in conducting these analyses:

1. Patients can be properly stratified into four CD4+ count categories.
2. All health care utilization, and therefore cost, is contingent upon CD4+ count at diagnosis.
3. Utilization patterns of patients identified through these programs can be modeled using utilization data from past HIV-positive patients treated at the same location.
4. All persons diagnosed are immediately linked to treatment.
5. All patients identified by Rapid Test would be diagnosed an average of 1.25 years later in the Usual Care program.

RESULTS

At all test offer rates and cohort sizes, there is no significant difference in total 7-year estimated costs of the Rapid Test and Usual Care programs (Table 3). At lower cohort sizes, the Rapid Test program is more costly than the Usual Care program, across all offer rates. As cohort size increases, there is a threshold at which Rapid Testing becomes less costly than Usual Care. The threshold is not consistent across offer rates due to the inclusion of implementation costs, which do not increase linearly with cohort size.

Implementation costs are responsible for the higher Rapid Testing costs at lower cohort sizes, as the organization is spending money in screening many patients but finding few who are HIV-positive. At lower cohort sizes, treatment costs are smaller; implementation costs thus represent a large percentage of overall costs. As cohort size—or number of patients identified as HIV-positive—increases, treatment costs increase substantially, whereas implementation costs increase only slightly. This is due to the fact that implementation costs are driven by test acceptance rates, whereas cohort size is substantially driven by disease prevalence. Hence, at large cohort sizes, when treatment costs are substantial in relation to implementation costs, Usual Care becomes the more costly program. Estimated costs for the Rapid Testing program differ across the same cohort size for the same reason. For example, the overall Rapid Testing costs for a cohort of size 11 may be $464,529 or $485,701 or $561,386. The reason for this discrepancy is implementation costs; the cost of offering tests to 10 patients in the ED is higher than the cost of offering HIV tests to 5 ED patients.

An examination of the relative cost components of each program reveals that the higher cost for Usual Care patients is largely due to more inpatient stays, reflecting more hospitalizations for these patients due to opportunistic infections (Table 4). Outpatient costs represent approximately one-fifth of costs for both programs, whereas pharmacy comprises the greatest percentage of costs for both programs. However, total pharmacy costs are lower for the Usual Care program, due to the fewer years of cost data.

As noted previously, the Usual Care program had no implementation costs, and only 5.75 years of data. This biases the cost estimates in its favor, as implementation costs for the Rapid Testing program range from $18,211 to $141,002. As statistical tests revealed no significant difference in the overall costs of the two programs, this suggests that once Usual Care patients are identified, their costs are higher than those for Rapid Test patients. Estimates of cost per capita reveal global costs per Usual Care patient to be $7237 annually vs. $5836 annually per Rapid Test patients.

DISCUSSION

Results indicate that over a 7-year period, an HIV screening program using a Rapid Testing approach is financially equivalent to following a Usual Care approach within the VA system. Given that early detection of HIV and linkage to treatment is also associated with better health outcomes, this analysis provides support for the implementation of a non-targeted oral Rapid Testing program.

Inpatient costs were substantially lower for Rapid Testing patients, offsetting the costs of implementing screening, and resulting in similar costs for the two HIV testing programs. Patients in the Usual Care program had higher inpatient costs, due to the fact that many of them were hospitalized for opportunistic infections (OIs). When modeling utilization of the Rapid Testing cohort, we excluded any hospitalization occurring within 3 months of HIV diagnosis. The model's exclusion of these inpatient costs for the Rapid Testing patients is partly responsible for the lower inpatient costs for this program. This exclusion was warranted, as we sought to quantify the marginal cost of implementing a non-targeted HIV Rapid Testing program. Including the inpatient costs of patients who were diagnosed due to hospitalization from OIs would misrepresent the clinical scenario of Rapid Testing patients, as patients who have acute OIs would have been diagnosed by Usual Care regardless of whether a Rapid Testing program was in place.

This analysis provides support for a non-targeted screening-based program. Although the initial investment of such a screening program may be as high as $141,000, the smaller treatment costs of patients identified at lower disease-severity offset the initial financial impact. This observation is crucial, as data from the VA and other health care systems indicate that targeted (risk-based) testing fails to capture a significant portion of the HIV-positive population (13,16).

This analysis also provides support for rapid testing. Although costs of rapid tests are higher than conventional assays, oral rapid tests are particularly attractive for use in ED settings due to higher patient acceptance rates and virtually immediate turnaround (5). Short turnaround increases the likelihood that patients will receive tests results, which is especially important in an ED-using population, where continuity of care can be problematic and many patients have difficulties returning for test results. Studies indicate that 21–30% of patients testing positive and 25–39% of patients testing negative for HIV do not return for the results of their tests (17,18).

This analysis used clinical and economic data from the VA. The advantage of using VA-specific data is that the VA is a non-profit, non-revenue-generating organization, which

ensures that financial data used represent actual costs, rather than profit or cost-shifting from uncompensated care. Use of actual data, rather than projections, serves to more accurately demonstrate the relationship between a program and its clinical and economic effects. Therefore, the results shown here are more indicative of the true cost of screening. Furthermore, this analysis demonstrates that the initial costs of ED-based screening are offset by lower inpatient utilization—data that are applicable to any integrated health system. Such results can guide other integrated systems in providing support for HIV screening in high-prevalence locations (defined as 0.5% or greater).

Limitations

The major limitation of this analysis lies in the small sample size used to determine utilization and estimate costs. This analysis had a sample size of 112, divided up into four disease-severity categories. First-order probabilistic sensitivity analyses were employed to gain more accurate estimates of the means with a small sample, but could not mitigate variation around the mean. Variation in inputs results in large confidence intervals, therefore increasing the likelihood of non-significance.

However, there were also important benefits from using these actual data. First, use of actual data takes into account the various drug (and therefore cost) combinations of HAART therapy, as well as the cost of modifying HAART therapy over time. Second, it better captures the relationship between patient drug use and hospitalization than modeling practices. Use of actual clinical data also takes into account that although all HIV-positive patients are referred to care, some did not actually enter care or cycled in and out of the system, and that rates of adherence to care vary considerably. Lastly, use of actual clinical data captures true physician practice patterns; using clinical practice guidelines to model care does not take into account physician deviation from these guidelines. Use of actual cost data, rather than charges or billing data, provides a more accurate indication of the real economic impact of each program.

In a clinical setting, it is likely that some patients would be lost to follow-up and not initiate treatment. However, this model assumed that all patients diagnosed were linked to treatment, as we believed that an organization should be prepared to allocate enough funding to a program such that all patients can be offered proper care for their disease.

This analysis was conducted using data from the VA system, which may limit generalizability of cost results to other settings. However, perhaps even more important than producing actual cost estimates, this analysis demonstrates the relationship between non-targeted screening and usual care—and provides evidence that a non-targeted screening program does not pose a significant cost burden to the organization. These findings are of use for decision-makers in any integrated system.

Treatment patterns for HIV may have changed over time, raising the issue of the appropriateness of using past data to model future utilization. Although new drugs may have come to market from FY 2000 to FY 2007, analyses of these data indicate that annual pharmacy costs have not changed over time. Inpatient costs have also not changed significantly, suggesting that even if drug therapy has altered, neither its costs nor its effect on inpatient utilization have been modified. Therefore, using past utilization data to inform this model was appropriate.

CONCLUSION

This article demonstrates the practical, real-world effect of implementing evidence-based policies, and adds a new component of support for HIV screening, by demonstrating that an

HIV screening program that utilizes rapid, non-targeted testing in an ED will not cost substantially more than a diagnostic testing approach.

Acknowledgments

This work was funded by the VA Health Services Research & Development Service Rapid Response Project (RRP) 07-283.

References

1. Glynn, M.; Rhodes, P. Estimated HIV prevalence in the United States at the end of 2003. National HIV Prevention Conference; Atlanta, GA. June 2005; Abstract T1-B1101. Available at: http://www. aegis.com/conferences/NHIVPC/2005/T1-B1101.html
2. Campsmith, M.; Rhode, P.; Hall, I. Estimated prevalence of undiagnosed HIV infection: US, end of 2006. 16th Conference on Retroviruses and Opportunistic Infections; Montreal, Canada. February 11, 2009; p. Abstract 1036
3. Samet JH, Freedberg KA, Savetsky JB, et al. Understanding delay to medical care for HIV infection: the long-term non-presenter. AIDS. 2001; 15:77–85. [PubMed: 11192871]
4. Sanders GD, Bayoumi AM, Sundaram V, et al. Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy. N Engl J Med. 2005; 352:570–85. [PubMed: 15703422]
5. Paltiel AD, Walensky RP, Schackman BR, et al. Expanded HIV screening in the United States: effect on clinical outcomes, HIV transmission, and costs. Ann Intern Med. 2006; 145:797–806. [PubMed: 17146064]
6. Paltiel AD, Weinstein MC, Kimmel AD, et al. Expanded screening for HIV in the United States—an analysis of cost-effectiveness. N Engl J Med. 2005; 352:586–95. [PubMed: 15703423]
7. Sanders GD, Bayoumi AM, Holodniy M, et al. Cost-effectiveness of HIV screening in patients older than 55 years of age. Ann Intern Med. 2008; 148:889–903. [PubMed: 18559840]
8. Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. MMWR Recomm Rep. 2006; 55:1–17. [PubMed: 16988643]
9. Qaseem A, Snow V, Shekelle P, et al. Screening for HIV in health care settings: a guidance statement from the American college of physicians and HIV medicine association. Ann Intern Med. 2009; 150:125–31. [PubMed: 19047022]
10. Mauskopf JA, Sullivan SD, Annemans L, et al. Principles of good practice for budget impact analysis: report of the ISPOR task force on good research practices- budget impact analysis. Value Health. 2007; 10:336–47. [PubMed: 17888098]
11. Shoenbaum EE, Webber MP, Vermund S, et al. HIV antibody in persons screened for syphilis: prevalence in a New York city emergency room and primary care clinics. Sex Transm Dis. 1990; 17:190–3. [PubMed: 2264007]
12. Goggin MA, Davidson AJ, Cantrill SV, et al. The extent of undiagnosed HIV infection among Emergency Department patients: results of a blinded seroprevalence survey and a pilot HIV testing program. J Emerg Med. 2000; 19:13–9. [PubMed: 10863112]
13. Patel AK, Patel KK. Future implications: compliance and failure with antriretroviral treatment. J Postgrad Med. 2006; 52:197–200. [PubMed: 16855321]
14. Fultz SL, Skanderson M, Mole LA, et al. Development and verification of a “virtual” cohort using the National VA Health Information System. Med Care. 2006; 44:S25–30. [PubMed: 16849965]
15. Cook J, Dasbach E, Coplan P, et al. Modeling the long-term outcomes and costs of HIV antiretroviral therapy using HIV RNA levels: application to a clinical trial. AIDS Res Hum Retroviruses. 1999; 15:499–508. [PubMed: 10221527]
16. Klein D, Hurley LB, Merrill D, Quesenberry CP Jr. Review of medical encounters in the 5 years before a diagnosis of HIV-1 infection: implications for early detection. J Acquir Immune Defic Syndr. 2003; 32:143–52. [PubMed: 12571523]

17. Owens DK, Sundaram V, Lazzeroni LC, et al. Prevalence of HIV infection among inpatients and outpatients in Department of Veterans Affairs health care systems: implications for screening programs for HIV. Am J Public Health. 2007; 97:2173–8. [PubMed: 17971545]

18. Centers for Disease Control and Prevention. HIV counseling with rapid tests. Oct 25. 2008 Available at: http://www.cdc.gov/hiv/topics/testing/resources/factsheets/print/rt_counseling.htm

Table 1

Cohort Sizes*

Offer Rate	Acceptance Rate	HIV Prevalence	Cohort Size
5/day	55%	0.5%	4
5/day	80%	0.5%	6
5/day	55%	1.0%	8
5/day	80%	1.0%	11
5/day	55%	3.0%	22
5/day	80%	3.0%	32
5/day	55%	5.4%	39
5/day	80%	5.4%	57
10/day	55%	0.5%	8
10/day	80%	0.5%	11
10/day	55%	1.0%	15
10/day	80%	1.0%	21
10/day	55%	3.0%	43
10/day	80%	3.0%	63
10/day	55%	5.4%	78
10/day	80%	5.4%	113
Universal	55%	0.5%	11
Universal	80%	0.5%	16
Universal	55%	1.0%	21
Universal	80%	1.0%	31
Universal	55%	3.0%	63
Universal	80%	3.0%	91
Universal	55%	5.4%	113
Universal	80%	5.4%	164

HIV = human immunodeficiency virus.

*Cohort size is dependent on four factors: the number of people visiting the Emergency Department annually, the percentage of patients offered testing, the percentage of patients accepting testing, and the prevalence of HIV. Due to uncertainty regarding the latter three rates, a variety of cohort sizes were modeled. Cohort size is calculated in the following manner: at an offer rate of 10 tests per business day, 2600 tests would be offered annually. If 55% of patients accept testing, 1430 tests would be conducted. At a 1% prevalence of disease, 14.3 patients would be identified as disease positive. Rounding up to the next whole person results in a cohort size of 15.

Table 2

Model Inputs: Percentage of Patients in Each CD4 Category

	CD4 < 50	CD4+ 50–199	CD4+ 200–350	CD4+ 350+
Usual Care	22%	26%	23%	29%
Rapid Testing	8%	24%	24%	43%

Table 3

Mean Overall Costs of Usual Care and Rapid Testing Programs in US Dollars*

Cohort Size	Offer Rate	Usual Care (95% CI)	Rapid Testing (95% CI)
4	5/day	131,143 (78,216–207,825)	175,114 (88,703–258,035)
6	5/day	244,115 (164,527–371,386)	262,056 (147,585–359,946)
8	5/day	311,373 (220,358–447,352)	332,389 (209,608–449,539)
11	5/day	449,675 (338,619–616,541)	464,529 (314,689–602,906)
22	5/day	918,463 (748,608–1,142,224)	916,894 (706,683–1,129,739)
32	5/day	1,373,520 (1,162,783–1,652,871)	1,334,239 (1,067,545–1,581,769)
39	5/day	1,661,119 (1,422,059–1,959,388)	1,638,520 (1,350,054–1,938,275)
57	5/day	2,448,055 (2,145,675–2,805,595)	2,383,208 (2,029,588–2,728,171)
8	10/day	311,373 (220,358–447,352)	348,036 (209,608–449,539)
11	10/day	449,675 (338,619–616,541)	485,701 (314,689–602,906)
15	10/day	630,549 (491,905–826,743)	653,836 (459,256–823,184)
21	10/day	878,091 (709,926–1,101,970)	899,234 (667,040–1,082,325)
43	10/day	1,840,167 (1,585,901–2,150,653)	1,828,215 (1,510,733–2,107,502)
63	10/day	2,692,457 (2,380,472–3,063,322)	2,679,336 (2,286,097–3,022,789)
78	10/day	3,372,705 (3,020,850–3,787,061)	3,287,460 (2,862,653–3,679,837)
113	10/day	4,874,310 (4,435,917–5,363,469)	4,773,087 (4,262,200–5,250,449)
11	Universal	449,675 (338,619–616,541)	561,386 (314,689–602,906)
16	Universal	671,123 (528,158–866,439)	785,252 (489,145–852,453)
21	Universal	878,091 (709,926–1,101,970)	976,253 (667,040–1,082,325)
31	Universal	1,320,338 (1,106,092–1,583,112)	1,418,088 (1,050,690–1,559,001)
63	Universal	2,692,457 (2,380,472–3,063,322)	2,756,595 (2,286,097–3,022,789)
91	Universal	3,936,820 (3,540,501–4,388,040)	3,927,737 (3,369,597–4,255,351)
113	Universal	4,874,310 (4,435,917–5,363,469)	4,848,802 (4,262,200–5,250,449)
164	Universal	7,103,144 (6,574,161–7,686,456)	7,011,975 (6,280,554–7,459,161)

*No differences were statistically significant.

Table 4

Average 7-year Cost Components

	Pharmacy	Inpatient	Outpatient
Usual Care	59%	22%	20%
Rapid Testing	64%	14%	21%

Substance Abuse, 33:227–230, 2012

ISSN: 0889-7077 print / 1547-0164 online
DOI: 10.1080/08897077.2011.640156

Routledge
Taylor & Francis Group

EDITORIAL

Screening, Brief Intervention, and Referral to Treatment (SBIRT) Curricular Innovations: Addressing a Training Gap

Unhealthy substance use, which includes the full spectrum of risky to dependent use of alcohol and other drugs, is common in patients cared for by health care providers in a variety of general health care settings. Early identification, assessment, treatment, and when indicated, referral to specialty care, of patients with unhealthy substance use can be challenging for both experienced health care providers as well as for health professional students and trainees. The clinical procedures of screening for unhealthy substance use, providing brief interventions to reduce unhealthy substance use, and referring for treatment patients in need of specialty treatment are evidence-based practices that are being promoted by research, policy, and service organizations in the United States. Despite this national movement, few health professional schools and programs have formal screening, brief intervention, and referral to treatment (SBIRT) curricula (1). This special issue is an early effort to describe and examine the outcomes of teaching SBIRT to health professionals and trainees.

Use of alcohol and illicit substances contributes to significant morbidity and mortality. Substance use has deleterious effects on common medical conditions (2–5), including diabetes mellitus, hypertension, sleep disorders, depression, chronic obstructive pulmonary disease, and osteoporosis (6–10). Moreover, unhealthy substance use increases risky behaviors, reduces access to preventive health care, reduces compliance with prescribed medical treatment, and adversely affects housing and employment (11–14). According to the 2009 National Survey on Drug Use and Health, of the almost 21 million people who needed treatment for illicit drug or alcohol use but did not receive it in 2008, 94% did not feel they needed treatment (15). The addiction specialist treatment system is not appropriate for all persons who have unhealthy alcohol use, or perhaps occasional substance use, nor can that system alone address the needs of all persons diagnosed with alcohol or substance use disorders. Other treatment modalities in a variety of clinical settings are needed.

SBIRT is an integrated, comprehensive, public health approach to addressing the full spectrum of unhealthy substance use in general health care settings. The core components of SBIRT include (1) screening—a strategy of early identification and assessment of individuals with unhealthy substance use through interview or self report; (2) brief intervention—a counseling approach that is focused on raising an individual's awareness of his or her unhealthy substance use and motivating that individual to a positive behavior change; and (3) referral to treatment—a proactive process that facilitates access to specialty addiction for individuals with substance use disorders.

In research studies, SBIRT for unhealthy alcohol use has demonstrated effectiveness, efficiency, and cost-effectiveness (16, 17). There is growing evidence for SBIRT for unhealthy drug use (18). Validated "single-item" screening questions make early identification of unhealthy substance use in busy general health care settings feasible (19, 20). Furthermore, models of implementation have been demonstrated for SBIRT in wide variety of general health care settings (21). SBIRT seems also to improve care coordination between general health care and specialty addiction services (22, 23). Yet despite evidence for the effectiveness of SBIRT, implementation of SBIRT into clinical settings has been slow (24, 25). An emerging body of literature aims to improve diffusion and application of SBIRT (26, 27). These guides promote interdisciplinary SBIRT planning and implementation teams, and additional calls have been made for the involvement of other health care professionals in SBIRT practices (28–30).

Although the Substance Abuse Mental Health Services Administration (SAMHSA), the Office of National Drug Control Policy (ONDCP), the United States Preventive Services Task Force (USPSTF), Institute of Medicine (IOM), and the Committee on Trauma of the American College of Surgeons (ACS) have promoted widespread SBIRT implementation, a variety of provider-, system-, and patient-level barriers to SBIRT delivery exist. These barriers include concerns about (1) lack of time and training for performing SBIRT, (2) perception or presence of more compelling clinical issues, (3) underutilization or lack of awareness of effective screening and treatment strategies, and (4) concerns about patient privacy and potential damage to the patient-provider relationship (31, 32). Additionally, providers may perceive that SBIRT for unhealthy substance use is simply ineffective, unsatisfying, uncomfortable, or not within their role responsibilities (26, 27).

SBIRT skills have been proposed as an addiction medicine core competency in generalist physician training (33). Recognizing the urgent need to train general health care providers in SBIRT, in 2008, the Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment (SAMHSA/CSAT) funded eleven 5-year SBIRT Medical Residency Program (SMRP) cooperative agreements to develop and implement training programs to "teach medical residents the skills to provide evidence-based SBIRT for patients who have, or are at risk for a substance use disorder [and] promote adoption of SBIRT through delivery of training to local and statewide medical communities for wider dissemination of SBIRT practices" (34).

Recognizing the critical need to disseminate curricular innovations regarding SBIRT training and evaluate the evidence of positive effects of SBIRT curricula on the learner- and patient-level outcomes, in January 2011, the Association of Medical Education and Research on Substance Abuse (AMERSA) led a call for papers of a special issue of the *Substance Abuse* journal. The special issue editors sought descriptions and evaluations of SBIRT curricula and training at health professional schools and residency and fellowship training programs, and targeting practicing health care professionals. Manuscripts describing innovative means to *develop*, *implement*, and *evaluate* SBIRT curricula were particularly sought. This special issue, "Screening and Brief Intervention for Substance Abuse," is the product of that call. In September 2011, we accepted 14 papers—4 full-length papers and 10 brief reports of curricular innovation—for publication.

The work contained in this special issue exemplifies early efforts to train health care professionals and their trainees in SBIRT skills. Skill proficiency of the learners in clinical practice is difficult to assess. Infusing SBIRT into academic curriculum can be challenging: the teachers of the curriculum often must be taught and making time for SBIRT knowledge and skill transfers may be difficult in the ever present setting of multiple competing curricular priorities. Pivotal questions such as who, when, and how to teach SBIRT to health care professionals need to be answered. For many of the educational projects described in this special issue, comprehensive outcome data have yet to be obtained. We will likely need to wait years to gain a full understanding of the effectiveness of various SBIRT curricula and training methods on changing health care providers' clinical practice. Ultimately

measuring SBIRT training on patient-level outcomes such as the number of patients receiving SBIRT services would be ideal. Yet, the work contained in the papers in this special issue reflect early attempts at describing curriculum development and implementation of addiction content into traditional learning environments. These educational program descriptions and early outcome assessments may serve as a template for diffusion of SBIRT curricula in other training environments. The authors of the work contained in this special issue should be commended for their pioneering SBIRT curriculum development and implementation efforts.

Adam J. Gordon, MD, MPH
VA Pittsburgh Healthcare System
and University of Pittsburgh
School of Medicine
Pittsburgh, PA

Daniel P. Alford, MD, MPH
Boston University School of Medicine
and Boston Medical Center
Boston, MA

REFERENCES

1. O'Connor PG, Nyquist JG, McLellan AT. Integrating addiction medicine into graduate medical education in primary care: the time has come. *Ann Intern Med.* 2011;154:56–59.

2. Gordon AJ, Gordon JM, Carl K, Hilton MT, Striebel J, Maher M. *Physical Illness and Drugs of Abuse: A Review of the Evidence*. Cambridge, UK: Cambridge University Press; 2010.

3. Brick J. *Handbook of the Medical Consequences of Alcohol and Drug Abuse.* 2nd ed. New York: The Haworth Press; 2008.

4. Frances RJ, Miller SI, Mack AH. *Clinical Textbook of Addictive Disorders. 3rd ed*. New York: The Guilford Press; 2005.

5. Rastegar DA, Fingerhood MI. *Addiction Medicine: An Evidence-Based Handbook.* Philadelphia: Lippincott Williams & Wilkins; 2005.

6. Howard AA, Arnsten JH, Gourevitch MN. Effect of alcohol consumption on diabetes mellitus: a systematic review. *Ann Intern Med.* 2004;140:211–219.

7. McFadden CB, Brensinger CM, Berlin JA, Townsend RR. Systematic review of the effect of daily alcohol intake on blood pressure. *Am J Hypertens.* 2005;18(2 Pt 1):276–286.

8. Mehra R, Moore BA, Crothers K, Tetrault J, Fiellin DA. The association between marijuana smoking and lung cancer: a systematic review. *Arch Intern Med.* 2006;166:1359–1367.

9. Stein MD, Herman DS, Bishop S, et al. Sleep disturbances among methadone maintained patients. *J Subst Abuse Treat.* 2004;26:175–180.

10. Sullivan LE, Fiellin DA, O'Connor PG. The prevalence and impact of alcohol problems in major depression: a systematic review. *Am J Med.* 2005;118:330–341.

11. Green TC, Kershaw T, Lin H, et al. Patterns of drug use and abuse among aging adults with and without HIV: a latent class analysis of a US Veteran cohort. *Drug Alcohol Depend.* 2010;110 (3):208–220.

12. Gordon AJ, Haas GL, Luther JF, Hilton MT, Goldstein G. Personal, medical, and healthcare utilization among homeless veterans served by metropolitan and nonmetropolitan veteran facilities. *Psychol Serv.* 2010;7:65–74.

13. Gordon AJ, Montlack ML, Freyder P, Johnson D, Bui T, Williams J. The Allegheny initiative for mental health integration for the homeless: integrating heterogeneous health services for homeless persons. *Am J Public Health.* 2007;97:401–405.

14. Lasser KE, Kim TW, Alford DP, et al. Is unhealthy substance use associated with failure to receive preventive services? *BMJ Open.* 2011; doi:10.1136/bmjopen-2010-000046.

15. Substance Abuse Mental Health Services Administration. *Results From the 2009 National Survey on Drug Use and Health: Volume 1. Summary of National Findings*. Rockville, MD: Office of Applied Studies; 2010.

16. United States Preventive Services Task Force. Screening and behavioral counseling interventions in primary care to reduce alcohol misuse: recommendation statement. *Ann Intern Med.* 2004;140:554–556.

17. Gentilello LM, Ebel BE, Wickizer TM, Salkever DS, Rivara FP. Alcohol interventions for trauma patients treated in emergency departments and hospitals: a cost benefit analysis. *Ann Surg.* 2005;241:541–550.

18. Bernstein J, Bernstein E, Tassiopoulos K, Heeren T, Levenson S, Hingson R. Brief motivational intervention at a clinic visit reduces cocaine and heroin use. *Drug Alcohol Depend.* 2005;77:49–59.

19. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A single-question screening test for drug use in primary care. *Arch Intern Med.* 2010;170:1155–1160.

20. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. Primary care validation of a single-question alcohol screening test. *J Gen Intern Med.* 2009;24:783–788.

21. Madras BK, Compton WM, Avula D, Stegbauer T, Stein JB, Clark HW. Screening, brief interventions, referral to treatment (SBIRT) for illicit drug and alcohol use at multiple healthcare sites: comparison at intake and 6 months later. *Drug Alcohol Depend.* 2009;99:280–295.

22. National Institute on Drug Abuse, NIH. Minutes of the 101st meeting of the National Advisory Council on Drug Abuse. Available at: http://www.drugabuse.gov/about/organization/nacda/minutes101st.html. Accessed April 1, 2011.

23. Alcohol Research & Health. Screening and Brief Intervention Part II—A Focus on Specific Settings. 2004/2005;28(2) Available at: http://pubs.niaaa.nih.gov/publications/arh28-2/toc28-2.htm. Accessed May 1, 2012.

24. Babor TF, Higgins-Biddle JC. Alcohol screening and brief intervention: dissemination strategies for medical practice and public health. *Addiction.* 2000;95:677–686.

25. Kuehn BM. Despite benefit, physicians slow to offer brief advice on harmful alcohol use. *JAMA.* 2008;299:751–753.

26. Desy PM, Perhats C. Alcohol screening, brief intervention, and referral in the emergency department: an implementation study. *J Emerg Nurs.* 2008;34:11–19.

27. Brooker C, Peters J, McCabe C, Short N. The views of nurses to the conduct of a randomised controlled trial of problem drinkers in an accident and emergency department. *Int J Nurs Stud.* 1999;36:33–39.

28. Anderson P, Aromaa S, Rosenbloom D, Enos G. *Screening & Brief Intervention: Making a Public Health Difference*. Boston, MA: Join Together, Boston University School of Public Health; 2008. Available at: http://www.adp.cahwnet.gov/SBI/pdfs/SBI_RPT_join_together.pdf. Accessed May 1, 2012.

29. Kaner E. Brief alcohol intervention: time for translational research. *Addiction.* 2010;105:960–961.

30. Nilsen P. Brief alcohol intervention—where to from here? Challenges remain for research and practice. *Addiction.* 2010;105:954–959.

31. Friedmann PD, McCullough D, Chin MH, Saitz R. Screening and intervention for alcohol problems. A national survey of primary care physicians and psychiatrists. *J Gen Intern Med.* 2000;15:84–91.

32. Saunders JB, Conigrave KM. Early identification of alcohol problems. *CMAJ.* 1990;143:1060–1069.

33. Jackson AH, Alford DP, Dube CE, Saitz R. Internal medicine residency training for unhealthy alcohol and other drug use: recommendations for curriculum design. *BMC Med Educ.* 2010;10:22.

34. United States Department of Health and Human Services, Substance Abuse & Mental Health Services Administration. SAMHSA Awards $66 Million for Programs Teaching Early Intervention Techniques to Use with Patients at Risk for Substance Abuse. Available at: http://www.samhsa.gov/newsroom/advisories/0809080442.aspx. Accessed May 1, 2012.

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

National patterns and predictors of liver biopsy use for management of hepatitis C

Erik J. Groessl[1,2,*], Lin Liu[2], Samuel B. Ho[1,2], Fasiha Kanwal[7,8], Allen L. Gifford[3,4], Steven M. Asch[5,6]

[1]*VA San Diego Healthcare System, USA;* [2]*University of California San Diego, USA;* [3]*Edith Nourse Rogers Memorial Veterans Hospital, USA;* [4]*Boston University, USA;* [5]*VA Greater Los Angeles Healthcare System, USA;* [6]*University of California Los Angeles, USA;* [7]*VA St. Louis, USA;* [8]*Washington University, USA*

Background & Aims: Liver biopsy remains the standard, recommended method for assessing liver damage associated with chronic hepatitis C (HCV) infection. However, there is considerable debate about how liver biopsy should best be used, especially with the advent of more efficacious antiviral therapies. To identify the factors that influence the use of liver biopsy for HCV patients, we describe variations in liver biopsy use at the delivery system and patient level in a national VA sample.

Methods: We analyzed VA HCV registry data for 171,893 VA patients with confirmed chronic HCV. Delivery system characteristics included geographic region and specialist time. Patient characteristics included antiviral treatment indicators, contraindications, volume of healthcare visits, and demographic variables. Logistic regression was used to explore correlates of biopsy use.

Results: Liver biopsy use in the VA system increased from 1997 to 2003 but began declining in 2004. Rates of liver biopsy from 2004 to 2006 varied by VA region, ranging from 5% to 18%. Treatment contraindications and laboratory tests were significantly associated with more biopsies. Demographic variables (higher age, lower BMI, race/ethnicity, and less% service connected disability) were associated with fewer biopsies. Regional variability remained significant independent of volume of care and specialist time.

Conclusions: Liver biopsy rates in the VA system have variability that seems unrelated to clinical need. New antiviral therapies and non-invasive assessment techniques may create additional uncertainty for the role of liver biopsy, perhaps explaining its decline in recent years. The availability of more effective antiviral therapies may also affect biopsy rates in the future.

Published by Elsevier B.V. on behalf of the European Association for the Study of the Liver.

Keywords: Hepatitis C; Liver biopsy; Guidelines.

Received 15 August 2011; received in revised form 16 February 2012; accepted 12 March 2012; available online 17 April 2012

* Corresponding author. Address: VA San Diego Healthcare System, 3550 La Jolla Village Drive, Mail Code 111 N-1 San Diego, CA 92161, USA. Tel.: +1 858 642 6347; fax: +1 858 552 4321.

E-mail address: egroessl@ucsd.edu (E.J. Groessl).

Abbreviations: HCV, hepatitis C virus; ALT, alanine aminotransferase.

Introduction

Hepatitis C virus (HCV) infects about 1.8% of the US population, [1] with higher rates in the VA patient population [2], and often leading to severe medical consequences [3–5]. Liver complications may be avoided with effective antiviral treatment. Antiviral treatment eliminated the virus in 40–50% of patients [6,7] until the recent availability of new therapies that have cure rates of 65–73% for genotype 1 patients in clinical trials [8,9]. Although liver biopsy is used for a variety of purposes, the most common usage is for the assessment of liver damage to assist with antiviral treatment decisions.

Liver biopsy is the recommended method for assessing liver damage resulting from chronic HCV infection. Both the VA [10] and NIH recommend its use for evaluating and making treatment decisions for most HCV patients (genotype 1). Additionally, the American Association of Liver Disease guidelines state that liver biopsy should be performed "when the results will influence antiviral treatment recommendations" [11,12].

Despite established recommendations, other factors such as alanine aminotransferase (ALT) elevation may influence biopsy rates. Some studies purport that HCV-infected persons with normal ALT levels have less liver damage than those with abnormal ALT levels [13], and that liver biopsy may be of little value for those with normal ALT [14]. Others believe all patients, regardless of ALT level, should be biopsied [15].

Relative contraindications to antiviral treatment such as depression or substance abuse can impact biopsy rates. Biopsy may only be considered for good treatment candidates. Conversely, biopsy results could be used to motivate lifestyle change to produce better treatment candidates. The literature on variation of health procedures in other conditions would suggest that non-clinical factors like race/ethnicity or geography may be another factor [16,17].

More recently, less invasive methods of liver fibrosis assessment have received considerable attention. A recent review of non-invasive methods concluded that non-invasive methods cannot match the information obtained from a biopsy [18]. Others have suggested that liver biopsy should only be recommended as a secondary test, due to the utility of biochemical marker tests [19].

Emerging HCV treatment options likely add to uncertainty about the utility of liver biopsy. New antiviral treatment may almost double sustained virologic response rates while shortening treatment duration [9,20,21]. With these improvements, liver biopsy may play a lesser role in treatment decision making.

Thus, there may be significant variability in how liver biopsy is used at different locations. In a recent clinician poll, substantial variability was found in recommendations for a hypothetical HCV patient. Of 3216 health professionals from 115 countries, 44% voted to perform a liver biopsy, 34% voted for expectant management with periodic assessment of liver function, and 22% voted to initiate HCV therapy with peginterferon and ribavirin [22]. Thus, despite the recommendations of expert bodies, the optimal use of liver biopsy in HCV-infected patients remains unclear.

Our study objective was to examine clinical and patient characteristics associated with variability of liver biopsy use for HCV-infected patients in VA Healthcare System.

Patients and methods

Design

The study was a retrospective, large database analysis using data from the VA Hepatitis C Clinical Case Registry from 1997 to 2006. Data analyses were conducted and refined from 2009 to 2011. Data analyzed for the study included patient demographics, outpatient visits, inpatient services, procedure codes, laboratory results, diagnosis codes, and problem lists. IRB approval and informed consent waiver were obtained.

Setting and sample

Research consisted only of retrospective database analyses. The original sample consisted of 304,024 VA patients who had been entered into the VA Hepatitis C Registry under unique patient identifiers from 1997 to 2006. Of those patients, 266,866 had evidence of either HCV antibody or confirmation testing. A subset of those, 171,893 patients, had clear evidence of a positive confirmatory test for chronic hepatitis C infection. Confirmed HCV positive was defined as any positive result from genotype, qualitative RNA and quantitative RNA. If a subject is HCV+ in a study year, the subject is HCV+ after this study year.

We used all available data (10 years) when examining the frequency of liver biopsies performed over time. To examine correlates and variability of liver biopsy use, we focused on the three most recent years of data (2004–2006) as most representative of current practice. "Receiving VA care" was defined as having one or more recorded inpatient or outpatient visits during this time period. This resulted in a sample size of 136,269 subjects who were confirmed to have chronic HCV infection and received VA care for chronic HCV from 2004 to 2006. Subjects with liver biopsies in any prior year or who died before 7/1/2004 were excluded.

Procedures

The initial data set included all VA patients with a positive HCV test. Since many patients with positive antibody tests clear the virus on their own, the first step was to limit our sample to patients that had tested positive on an HCV confirmation test. Confirmatory testing was defined as a positive result on either a qualitative or quantitative HCV RNA test, or a positive HCV genotype test. Confirmatory tests with 76 different name variations were classified into one of the three types of confirmatory tests. Next, to reduce the chance of including test results that were labeled incorrectly, test result values were examined to ensure consistency with the type of test (qualitative test – yes/no answers, quantitative tests – numerical values within range, genotype tests – ranging from 1 to 6). Approximately 5% of the results were indeterminable or ambiguous and could not be used for analysis.

We defined receipt of liver biopsy as presence of the following codes in the outpatient or inpatient data files (50.11 – Closed Liver Biopsy, 50.12 or 50.121 – Open Liver Biopsy, 50.13 – Transjugular Liver Biopsy, 50.14 – Laparoscopic Liver Biopsy, 47000 – Needle Biopsy of Liver, 47001 – Needle Biopsy, Liver Add On, 47100 – Wedge Biopsy of Liver).

We then examined the variables associated with receipt of a liver biopsy. To identify predictors of liver biopsy, we began with a list of possible contraindications for antiviral treatment developed by Kanwal *et al.* for VA patients with HCV [23]. Possible contraindications were determined by ICD-9 codes and not actual treatment decisions. Because the goals of our study differed from the study cited, modifications were made to the coding of these contraindications. One modification was to differentiate current substance use and psychiatric disorders from long-term disorders because long-term or permanent disorders may not be a current problem or contraindication by themselves, but may be related to biopsy rates. A substance use or psychiatric disorder was considered "current" if any inpatient or outpatient visits between 2003 and 2006 had substance use or psychiatric ICD 9 codes. Substance use or psychiatric disorders that appeared on the medical record "problem list", but were not linked to visits between 2003 and 2006, were considered long-term. A second modification from the reference manuscript [23] was that laboratory data was not used in the algorithms for defining the presence of co-morbid conditions as contraindications. Laboratory variables were examined directly as independent correlates of receipt of liver biopsy (Table 1).

Demographic variables were recoded into the categories indicated in Table 3. Regional variation was examined using the 21 regions of the national VA healthcare system. To examine the effect of specialist time on biopsy frequency, we obtained data from the 2007 Survey to Assess Hepatitis C Care in VHA conducted by the VA Public Health Strategic Health Care Group. Responses from 141 "lead HCV clinicians" at VHA facilities indicated the "Number of weekly ½ days clinics covered" by 13 different types of healthcare providers. To account for differences in the number of patients served by each region, we divided the average number of ½ day clinics for hepatologists and gastroenterologists in each region by the number of HCV patients from that region.

Statistical analyses

We present raw numbers of biopsies over time and proportion of patients in the database who were biopsied each year between 1997 and 2006. We focused on raw numbers instead of the proportion because many new patients were added to the registry each year, making the proportion drop each year while the number of biopsies being conducted was increasing. To examine geographical variability, descriptive statistics and Chi-square analysis were used to describe and examine differences in the proportion of patients biopsied at each of the 21 VA healthcare regions between 2004 and 2006. We limited the analysis of regional variation to the most recent years to best represent current utilization.

Multivariate logistic regression was used to examine the effect of patient demographic and clinical characteristics, and geographic regions on the receipt of liver biopsy in HCV-infected patients. Because liver biopsy may be used for different reasons among distinct subgroups within our sample, we identified three patient subgroups (liver transplant, cirrhosis prior to biopsy, and those without transplant or cirrhosis prior to biopsy). Our final model (Table 3) focuses on HCV-infected patients without cirrhosis or liver transplant, representing over 90% of the sample.

Independent variables were first examined in bivariate analyses with presence of liver biopsy between 2004 and 2006 as the dependent variable. Variables with p-value <0.15 in bivariate analyses were included as predictors in multivariate logistic regression. Variables were entered into the final model in the following order: contraindications, labs, demographics and other variables (specialist time, geographic region). Stepwise backward model selection for main effects in the final model was performed. Variables for the initial model were selected using Akaike's Information Criterion. Then the variable with the largest p-value was removed from the model until all variables included in the model with p-values <0.05 were examined. Non-parametric rank sum tests confirmed that specialist time differed by geographic region. Analyses were performed with SAS (version 9.1) and the open source statistical package R (version 2.8.1).

Results

Rates of liver biopsy and patient characteristics

The raw number of liver biopsies conducted in the VA system nationally in 1997 was only 475. This number grew quickly in subsequent years and peaked at 4758 in 2004 before declining slightly in 2005 and 2006 (Fig. 1). As an overall proportion for the study period of 1997–2006, 16.7% (28,677 of 171,893) of VA patients with confirmed HCV infection received a biopsy.

Research Article

Table 1. Contraindication and laboratory variables.

Contraindications	ICD9 codes
Major, uncontrolled depressive illness	296.2x-3x (1 inpatient or 2 outpatient)
Current illicit drug/alcohol use	305.xx except 305.1 (1 inpatient or 2 outpatient)
Untreated hyperthyroidism	242.xx (1 inpatient or 1 outpatient)
Severe hypertension	401.xx-405.xx (1 inpatient or 1 outpatient)
Severe heart failure	428.xx (1 inpatient or 2 outpatient)
Significant coronary artery disease	414.xx (1 inpatient or 2 outpatient)
Poorly controlled diabetes	250.xx (2 inpatient or 2 outpatient)
Severe COPD	490.xx-496.xx (1 inpatient)
Severe psychiatric disease	300.xx, 309.81, 295.xx-298.xx (1 inpatient)
HIV co-infection	V08, 042 (1 inpatient or 1 outpatient)
Chronic renal disease	585.xx (1 inpatient or 1 outpatient)
Decompensated cirrhosis	571.2, 571.5-6 (1 inpatient or 1 outpatient)
Liver transplant	V42.7 (1 inpatient or 1 outpatient)
Renal, heart, and lung transplant	V42.0-1, V42.6, V58.44, E878.0 (1 inpatient or 1 outpatient)
Laboratory variables*	**Variable coding**
Hepatitis A	has been tested (Yes/No)
Hepatitis B	has been tested (Yes/No)
HCV genotype	1, 2, 3, or 4
Total bilirubin	high abnormal (Yes/No)
Aspartate aminotransferase (AST)	high abnormal (Yes/No)
Alanine aminotransferase (ALT)	high abnormal (Yes/No)
Platelet count	low abnormal (Yes/No)
Albumin	low abnormal (Yes/No)
Hemoglobin	low abnormal (Yes/No)
White blood count (WBC)	low abnormal (Yes/No)
Creatinine	high abnormal (Yes/No)

*For subjects with a liver biopsy from 2004 and 2006, laboratory tests between 1/1/2003 and the date of liver biopsy were used. For subjects without a liver biopsy from 2004 and 2006, laboratory tests conducted between 1/1/2003 and 09/30/2006 were used.

However, many of these individuals may have had contraindications for antiviral treatment, reducing the need for biopsy. Clinical and socio-demographic characteristics of the sample from 2004 to 2006 were compared for patients who did and did not receive a liver biopsy (Table 2). Given the large sample size, almost all variables were significantly different at the p <0.001 level.

Geographical variability in biopsy rates

To study geographical variability, we examined the proportion of HCV-infected patients who were biopsied by region from 2004 to 2006. The overall rate of biopsy among HCV patients in care from 2004 to 2006 was 9.3%. Overall data include many patients who had likely not been biopsied for specific reasons or refused (Fig. 2). When broken down by clinical subgroup, rates were 19.9% among liver transplant recipients (0.8% of the sample), 9.8% among HCV patients with cirrhosis (9.1% of the sample), and 9.2% among those without transplant or cirrhosis (90.1% of the sample). In initial models, patients with transplant (OR 3.29, <0.001) and cirrhosis prior to biopsy (OR 1.20, <0.001) were significantly more likely to be biopsied.

There was significant variation in the rates of liver biopsy across the 21 VA regions. While 17.8% of HCV patients in region U received a liver biopsy between 2004 and 2006, biopsy rates in regions G, H, and P were much lower, in the range of 5–7% (p <0.001). Removing all patients with a potential contraindication, increased the proportion of patients biopsied from 9.4% to 12.1% during that period. Regions B and U biopsied over 21% of patients without contraindications during this period while other regions were as low as 6.6%.

Factors associated with receipt of liver biopsy

Most contraindications were significantly associated with lower rates of liver biopsy in the final model. Recent organ transplant and advanced liver disease were significantly associated with higher rates of liver biopsy (Table 2).

As expected, genotype 2 and 3 patients were less likely than genotype 1 patients to receive a biopsy. Patients who had been tested for hepatitis A or B were more likely to be biopsied. HIV infection was not significantly associated with liver biopsy. Liver biopsy rates decreased slightly with each year of age, and increased slightly with increasing BMI and service connection

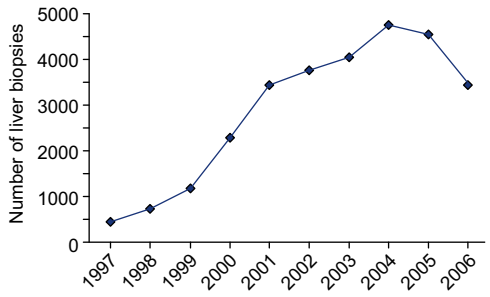


Fig. 1. Number of liver biopsies by year, national data.

percentage. When compared with non-Hispanic Whites, Hispanics, African-Americans, and patients of Other/Unknown race/ethnicity were significantly less likely to receive a biopsy.

In the next regression step, hepatologist and gastroenterologist specialist times were initially associated with more biopsies, but became non-significant when the mean number of healthcare visits (associated more liver biopsies) was entered into the analysis. Finally, the 21 categories of geographic VA region were entered while controlling for other variables in the model. Specialist time remained significant after geographic region was entered, indicating independent effects. With region A's biopsy rate of 13% as the reference group, only regions B, E, F, and M, were not significantly different. Fourteen regions had significantly lower biopsy rates while subject two regions (O, U) had higher rates.

To examine the univariate connection between liver biopsy rates and antiviral treatment rates, we examined correlation of the two rates by geographic regions. A significant correlation was found between liver biopsy rates and HCV antiviral treatment rates ($r(21) = 0.46$, $p = 0.036$).

Discussion

The burden of chronic hepatitis C continues to grow in the US, especially for the VA healthcare system, which treats more HCV-infected individuals than any other system. Liver biopsy has long been the gold standard for assessing the severity of liver damage among HCV-infected patients, and it is recommended for the evaluation of all HCV patients with genotype 1 virus [10], who comprise about 70% of all cases in the US. Despite these recommendations, our results indicate that only about 17% of all HCV-infected patients receiving care in the VA system had been biopsied.

There are a number of plausible clinical and policy-related reasons that doctors and patients may choose to forgo biopsy. One common reason is that the majority of VA patients are not immediate candidates for antiviral medications because of relative contraindications, most commonly substance use and psychological disorders [24]. Because liver biopsy can have side effects, result in complications, and consume healthcare resources, the benefits of a biopsy may or may not outweigh the "costs" when being considered by healthcare providers and their patients. In addition, motivated candidates for antiviral therapy may decide to start therapy regardless of the additional prognostic information they might get from biopsy, and thus might forgo it.

In our large national sample, liver biopsy use increased between 1997 and 2004 but began to decline in 2005 and 2006, despite the growing cumulative number of VA patients who have been confirmed with chronic HCV. Although the cumulative number of unbiopsied patients was increasing during the studied time period, these patients may have decided in conjunction with their providers to not get biopsied, so a majority of the most appropriate candidates for biopsy may have already received one. As laboratory and other non-invasive assessments of liver fibrosis become more accurate [25] and are recommended liver disease guidelines [26], some reduction in the need for liver biopsy may be expected [27].

Other possible factors that may influence the consistency of liver biopsy use in the VA Healthcare system include: (a) the emergence of less invasive techniques for assessing liver damage (b) the belief that all HCV patients should be treated; and (c) the ongoing development of new antiviral medications with fewer side effects and high SVR rates.

An interesting finding among analyses exploring predictors of biopsy was that Hispanics, African-Americans, and patients of Other/Unknown race/ethnicity were less likely to receive a biopsy than non-Hispanic Whites. It is important to note that the rates of biopsy in African-Americans were only slightly lower than those of non-Hispanic Whites, but were statistically significant because of the very large sample size. Future research is needed to explore whether these differences might be explained by cultural preferences, access to care, or provider behavior in connection with rates of antiviral treatment. Understanding the cultural preferences and access to care among these groups may improve the quality of care they receive.

Results also indicate that there is considerable variability between geographic regions in the proportion of HCV patients receiving biopsies. This is not surprising as variation in the rates of medical procedures and services has been well documented [28,29]. The variability patterns for biopsy in the VA systems are similar to the variability patterns in rates of antiviral treatment in the VA system [24,30], suggesting that higher rates of biopsy are associated with higher rates of treatment. Hepatologist and gastroenterologist specialist times were not associated with more liver biopsies once the overall volume of health care received entered the model as a significant co-variate. Notably, regional variability remained significant after controlling for clinical characteristics, patient demographics, or specialist time. Specialist time or overall volume of care has frequently affected procedure variability in other conditions, but does not always fully explain geographic differences [31–34]. Other explanations for regional variability in liver biopsies are not clear, but some evidence suggests that region U had the highest rates of biopsy and antiviral treatment because of a particular HCV care model developed at the main VA Medical Center in that region [35,36]. Other possible reasons for geographic variability in biopsy rates are differences in available resources, varying definitions of antiviral treatment candidacy, and varying beliefs about the utility of biopsy associated with proximity to large academic medical centers.

Our study confirms and strengthens the data above in demonstrating geographic differences in performance of a common and expensive diagnostic practice in a single uniform medical care system. Because it is within a single-payer and single administrative system, variables such as differences in market forces, financial incentives, and Medicare reimbursement policy

Research Article

Table 2. Patient characteristics and univariate relationships with liver biopsy.

Variables	Biopsy n = 12,753 % or mean (SD)	No biopsy n = 123,516 % or mean (SD)	*p* value
Socio-demographic characteristics			
Age (yr)	51.5 (5.8)	52.8 (7.9)	<0.001
BMI	28.0 (5.2)	27.5 (5.4)	<0.001
% Service connected	19.8 (32.9)	19.3 (33.3)	<0.001
Race/ethnicity (ref. group White)			
Non-Hispanic White	45.4%	37.7%	<0.001
Hispanic White	3.5%	4.0%	
Native American	0.4%	0.5%	
Asian Pacific Islander	0.8%	0.7%	
African American	28.2%	27.5%	
Others/Unknown	21.8%	29.6%	
Long-term psychiatric disorder	28.4%	35.8%	<0.001
Long-term substance use disorders	25.4%	34.4%	<0.001
No. of HCV-related visits/month	1.73	1.37	<0.001
Laboratory data*			
ALT (high abnormal)	3.34 (4.7)	3.28 (4.9)	<0.001
AST (high abnormal)	3.5 (5.4)	3.9 (6.0)	<0.001
White blood count (low abnormal)	1.2 (4.4)	1.7 (5.5)	<0.001
Creatinine (high abnormal)	0.9 (5.4)	1.7 (9.2)	<0.001
Bilirubin (high abnormal)	0.9 (4.1)	1.1 (4.0)	<0.001
Platelet count (low abnormal)	1.6 (6.7)	2.4 (8.1)	<0.001
Albumin (low abnormal)	1.4 (5.5)	2.1 (6.4)	<0.001
Hemoglobin (low abnormal)	2.3 (8.9)	4.3 (13.0)	<0.001
Hep A vaccinated	53.0%	40.0%	<0.001
Hep B vaccinated	67.4%	54.5%	<0.001
Genotype			<0.001
1	82%	79.5%	
2	10.1%	12.2	
3	7.0%	7.5	
4	1.0%	0.8	
Contraindications			
Active substance use disorder	22.6%	31.4%	<0.001
Major depression	9.4%	11.3%	<0.001
Other severe psychiatric disorder	8.8%	14.0%	<0.001
HIV	3.4%	4.1%	<0.001
Transplant - renal, heart, or lung	0.5	0.4	0.058
Thyroid disorder	0.5	0.7	0.004
Significant hypertension	48.9%	55.8%	<0.001
Severe heart failure	1.3%	4.1%	<0.001
Coronary artery disease	5.2%	9.5%	<0.001
Diabetes	17.0%	20.9%	<0.001
Chronic obstructive pulmonary disease	4.1%	8.1%	<0.001
Chronic renal disease	2.2%	4.4%	<0.001
Cirrhosis	10.7%	9.5%	<0.001
Liver transplant	1.7%	0.7%	<0.001

*Lab tests data shown are the mean number of abnormal (high or low) lab tests during the study period.

differences do not apply. In addition, because patient socio-demographic and health-related diagnoses were taken directly from the national computerized medical record, it is unlikely that important patient characteristics and diagnoses were missed or not taken into account. The present data strongly support the conclusion that geographic factors are associated with variation in biopsy use, independent of patient-related and administrative factors.

Table 3. Logistic regression model for liver biopsy in patients without transplant or cirrhosis.

	Odds ratio	*p* value
Contraindications		
Transplant-non-liver	1.328	0.255
Thyroid	0.740	0.042
Hypertension	0.795	<0.001
Heart	0.578	<0.001
Coronary	0.717	<0.001
Diabetes	0.791	<0.001
Pulmonary	0.649	<0.001
Active substance use	0.632	<0.001
Renal	0.623	<0.001
Current psychiatric disorder	0.720	<0.001
Laboratory data (normal *vs.* abnormal)		
ALT	1.000	0.959
AST	0.965	<0.001
HG	0.969	<0.001
WBC	0.931	<0.001
Creatinine	1.008	<0.039
Hep A	1.283	<0.001
Hep B	1.182	<0.001
Genotype (ref. group genotype 1)		
2	0.739	<0.001
3	0.813	<0.001
4	1.169	0.184
Sociodemographic characteristics		
Age	0.990	<0.001
BMI	1.011	<0.001
Race/ethnicity (ref. group White)		
Hispanic White	0.789	<0.001
Native American	0.664	0.024
Asian Pacific islander	0.881	0.318
African American	0.941	0.042
Others/Unknown	0.589	<0.001
Service connection	1.003	<0.001
History of psychiatric problems	0.685	<0.001
History of substance use disorders	0.613	<0.001
No. of HCV-related visits	1.187	<0.001
Regional characteristics		
Hepatologist time in HCV clinic	1.009	0.233
Gastroenterologist time in HCV clinic	0.995	0.629
VA region (ref. group region A)		
B	0.991	0.914
C	0.643	<0.001
D	0.330	<0.001
E	0.857	0.047
F	0.999	0.992
G	0.392	<0.001
H	0.546	<0.001
I	0.550	<0.001
J	0.508	<0.001
K	0.645	<0.001
L	0.751	<0.001
M	1.027	0.720
N	0.542	<0.001
O	1.663	<0.001
P	0.412	<0.001
Q	0.371	<0.001
R	0.850	0.027
S	0.752	<0.001
T	0.550	<0.001
U	2.149	<0.001

Although our results show that liver biopsy has been used inconsistently in the recent past, and many genotype 1 HCV-infected individuals have never received a biopsy, it is hard to determine the optimal rate of biopsy and the impact of these results on the quality of care provided. Liver biopsy correlates fairly well with antiviral treatment rates and is recommended by the VA and other organizations as an important step in antiviral treatment decision-making for genotype 1 patients [10,11,26]. A recent study on the quality of care in the VA system calls for higher treatment rates and efforts to improve SVR [37]. Thus, if the recommendations noted above are followed, liver biopsy use should increase as treatment rates increase. Although it remains possible that the higher SVR rates of the new antiviral therapies may lead to less reliance on liver biopsy as it does with genotype 2–3 patients, biopsy will likely retain an important role because of the need to prioritize patients for treatment, given available resources. Therefore, it seems important for VA and non-VA clinicians and researchers to study this use and provide more detailed guidance on the use of liver biopsy with HCV patients in the near future. The procedure will likely retain its value in some circumstances, but its utility must be weighed against its cost and potential for infrequent but significant side effects. Patient preferences and behaviors should also be considered when gauging optimal rates of liver biopsy.

A limitation of the data is that it relies upon the accuracy of the VA Hepatitis C Registry. Some lab tests may be mislabeled, and mistakes likely occur in the human data entry process. It is impossible to know the exact level of accuracy, but in general, after recoding obvious transpositional errors, we excluded data on only a couple thousand cases, constituting 1–2% of cases for overall analyses.

Our results include all available data from the VA Hepatitis C Registry. Although some data show that treatment rates and characteristics of patients with HCV do not differ much between

Research Article

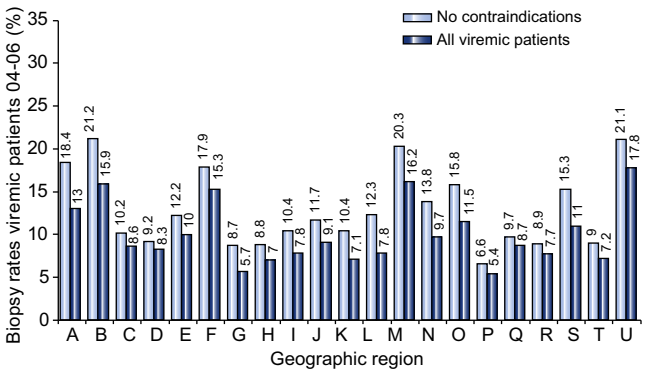


Fig. 2. Rates of liver biopsy among viremic patients 2004–2006.

the VA and community clinics [24], conclusions from our results should only be applied to the VA system and may not generalize to non-VA settings.

In conclusion, we found that many VA patients with HCV have not been biopsied and the number of liver biopsies being conducted within the VA system nationally has been declining despite a large pool of untreated patients with HCV. We also found that there is considerable variability in the rate at which liver biopsy is used among the 21 regions in the VA system and to a lesser extent by other factors seemingly unrelated to clinical need. Future research and policy should focus on developing more detailed guidance on the use of liver biopsy among HCV-infected patients, as non-invasive techniques develop and in light of the recent arrival of more efficacious antiviral treatments.

Financial support

VA HSR&D, Project SHP – 08–203.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Acknowledgements

This study was funded by a research grant from the VA HSR&D, Project SHP – 08-203.

References

[1] Alter MJ, Kruszon-Moran D, Nainan OV, McQuillan GM, Gao F, Moyer LA, et al. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. N Engl J Med 1999;341:556–562.

[2] Dominitz J, Boyko EJ, Koepsell TD, Heagerty P, Maynard C, Sporleder JL. Prevalence of hepatitis C in veterans: results of VA cooperative study #488. Gastroenterology 2004;126:A672.

[3] CDC. Center for Disease Control: recommendations for prevention and control of HCV infection and HCV-related chronic disease. Morbidity Mortality Weekly Rep 1998;47:1–10.

[4] CDC CfDC. Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. Morbidity Mortality Weekly Rep 1998;47:1–39.

[5] Heintges T, Wands JR. Hepatitis C virus: epidemiology and transmission. Hepatology 1997;26:521–526.

[6] Fried MW, Hadziyannis SJ. Treatment of chronic hepatitis C infection with peginterferons plus ribavirin. Semin Liver Dis 2004;24 Suppl 2:47–54.

[7] Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales Jr FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 2002;347:975–982.

[8] Burney T, Dusheiko G. Overview of the PROVE studies evaluating the use of telaprevir in chronic hepatitis C genotype 1 patients. Expert Rev Anti-infect Ther 2011;9:151–160.

[9] Poordad F, McCone Jr J, Bacon BR, Bruno S, Manns MP, Sulkowski MS, et al. Boceprevir for untreated chronic HCV genotype 1 infection. N Engl J Med 2011;364:1195–1206.

[10] Yee HS, Currie SL, Darling JM, Wright TL. Management and treatment of hepatitis C viral infection: recommendations from the department of veterans affairs hepatitis C resource center program and the national hepatitis C program office. Am J Gastroenterol 2006;101:2360–2378.

[11] Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 2009;49:1335–1374.

[12] Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C. Hepatology 2004;39:1147–1171.

[13] Persico M, Perrotta S, Persico E, Terracciano L, Folgori A, Ruggeri L, et al. Hepatitis C virus carriers with persistently normal ALT levels: biological peculiarities and update of the natural history of liver disease at 10 years. J Viral Hepat 2006;13:290–296.

[14] Pasquale G, Sagnelli E, Coppola N, Scarano F, Scolastico C, Bellomo PF, et al. Is liver biopsy necessary for hepatitis C virus carriers with persistently normal aminotransferase levels? Eur J Gastroenterol Hepatol 2003;15:831–833.

[15] Puoti C, Bellis L, Guarisco R. Liver biopsy in hepatitis C virus carriers with normal alanine aminotransferase levels. Hepatology 2009;50:655.

[16] Baicker K, Chandra A, Skinner JS, Wennberg JE. Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries. Health Aff (Millwood) 2004;Suppl. Web Exclusives:VAR33–VAR44.

[17] Skinner J, Weinstein JN, Sporer SM, Wennberg JE. Racial, ethnic, and geographic disparities in rates of knee arthroplasty among medicare patients. N Engl J Med 2003;349:1350–1359.

[18] Smith JO, Sterling RK. Systematic review: non-invasive methods of fibrosis analysis in chronic hepatitis C. Aliment Pharmacol Ther 2009;30:557–576.

[19] Poynard T, Imbert-Bismut F, Munteanu M, Messous D, Myers RP, Thabut D, et al. Overview of the diagnostic value of biochemical markers of liver fibrosis (FibroTest, HCV FibroSure) and necrosis (ActiTest) in patients with chronic hepatitis C. Comp Hepatol 2004;3:8.

[20] Kronenberger B, Zeuzem S. Current and future treatment options for HCV. Ann Hepatol 2009;8:103–112.

[21] McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, Kauffman R, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. N Engl J Med 2009;360:1827–1838.

[22] Wilck MB, Hamel MB, Baden LR. Clinical decisions. Management of incidental hepatitis C virus infection – polling results. N Engl J Med 2009;360:e30.

[23] Kanwal F, Hoang T, Spiegel BM, Eisen S, Dominitz JA, Gifford A, et al. Predictors of treatment in patients with chronic hepatitis C infection – role of patient versus nonpatient factors. Hepatology 2007;46:1741–1749.

[24] Ho SB, Groessl E, Dollarhide A, Robinson S, Kravetz D, Dieperink E. Management of chronic hepatitis C in veterans: the potential of integrated care models. Am J Gastroenterol 2008;103:1810–1823.

[25] Usluer G, Erben N, Aykin N, Dagli O, Aydogdu O, Barut S, et al. Comparison of non-invasive fibrosis markers and classical liver biopsy in chronic hepatitis C. Eur J Clin Microbiol Infect Dis 2012 [Epub ahead of print].

[26] Calvaruso V, Craxi A. Of the study of the liver hepatitis C virus clinical practice guidelines. Liver Int 2012;32 (Suppl. 1):2–8.

[27] Lai M, Afdhal NH. Editorial: staging liver fibrosis in hepatitis C: a challenge for this decade. Am J Gastroenterol 2011;106:2121–2122.

[28] Sutherland JM, Fisher ES, Skinner JS. Getting past denial – the high cost of health care in the United States. N Engl J Med 2009;361:1227–1230.

[29] Wennberg JE. Practice variation: implications for our health care system. Managed care (Langhorne, Pa) 2004;13:3–7.

[30] Backus LI, Boothroyd DB, Phillips BR, Mole LA. Pretreatment assessment and predictors of hepatitis C virus treatment in US veterans coinfected with HIV and hepatitis C virus. J Viral Hepat 2006;13:799–810.

[31] Birkmeyer JD, Sharp SM, Finlayson SR, Fisher ES, Wennberg JE. Variation profiles of common surgical procedures. Surgery 1998;124:917–923.

[32] Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. N Engl J Med 2010;363:45–53.

[33] Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ 2002;325:961–964.
[34] Wennberg JE. Practice variations and health care reform: connecting the dots. Health Aff (Millwood) 2004;Suppl. Web Exclusives:VAR140–VAR144.
[35] Groom H, Dieperink E, Nelson DB, Garrard J, Johnson JR, Ewing SL, et al. Outcomes of a hepatitis C screening program at a large urban VA medical center. J Clin Gastroenterol 2008;42:97–106.
[36] Ho SB, Gebhard RL, Tetrick LL, Durfee JM, Steinberg N, Parenti CM, et al. Key components and logistics of hepatitis C care within a VISN: a proposal. Veterans Health Syst J 1999;4:41–55.
[37] Kramer JR, Kanwal F, Richardson P, Mei M, El-Serag HB. Gaps in the achievement of effectiveness of HCV treatment in national VA practice. J Hepatol 2012;56:320–325.

Effects of Alcohol Retail Privatization on Excessive Alcohol Consumption and Related Harms

A Community Guide Systematic Review

Robert A. Hahn, PhD, MPH, Jennifer Cook Middleton, PhD, Randy Elder, PhD, Robert Brewer, MD, MSPH, Jonathan Fielding, MD, MPH, MBA, Timothy S. Naimi, MD, MPH, Traci L. Toomey, PhD, Sajal Chattopadhyay, PhD, Briana Lawrence, MPH, Carla Alexia Campbell, MHSc, and the Community Preventive Services Task Force

Context: Excessive alcohol consumption is the third-leading cause of preventable death in the U.S. This systematic review is one in a series exploring effectiveness of interventions to reduce alcohol-related harms.

Evidence acquisition: The focus of this review was on studies evaluating the effects of the privatization of alcohol retail sales on excessive alcohol consumption and related harms. Using *Community Guide* methods for conducting systematic reviews, a systematic search was conducted in multiple databases up to December 2010. Reference lists of acquired articles and review papers were also scanned for additional studies.

Evidence synthesis: A total of 17 studies assessed the impact of privatizing retail alcohol sales on the per capita alcohol consumption, a well-established proxy for excessive alcohol consumption; 9 of these studies also examined the effects of privatization on the per capita consumption of alcoholic beverages that were not privatized. One cohort study in Finland assessed the impact of privatizing the sales of medium-strength beer (MSB) on self-reported alcohol consumption. One study in Sweden assessed the impact of re-monopolizing the sale of MSB on alcohol-related harms. Across the 17 studies, there was a 44.4% median increase in the per capita sales of privatized beverages in locations that privatized retail alcohol sales (interquartile interval: 4.5% to 122.5%). During the same time period, sales of nonprivatized alcoholic beverages decreased by a median of 2.2% (interquartile interval: −6.6% to −0.1%). Privatizing the sale of MSB in Finland was associated with a mean increase in alcohol consumption of 1.7 liters of pure alcohol per person per year. Re-monopolization of the sale of MSB in Sweden was associated with a general reduction in alcohol-related harms.

Conclusions: According to *Community Guide* rules of evidence, there is strong evidence that privatization of retail alcohol sales leads to increases in excessive alcohol consumption.

(Am J Prev Med 2012;42(4):418–427) Published by Elsevier Inc. on behalf of American Journal of Preventive Medicine

Context

Excessive alcohol consumption, including both binge drinking and underage drinking, is responsible for approximately 79,000 deaths per year in the U.S., making it the third-leading cause of preventable death in the nation.[1] In 2009, approximately 23% of adult drinkers (aged ≥18 years) in the U.S. reported binge drinking (consuming five or more drinks per occasion for men and four or more drinks per occasion for women) in the past 30 days, as did 25.2% of high school students.[2,3] Among full-time college students in 2008, 48.6% of men and 34.4% of women reported binge drinking.[4] In 2006,

From the Community Guide Branch, Epidemiology and Analysis Program Office (Hahn, Middleton, Elder, Chattopadhyay, Lawrence, Campbell), National Center for Chronic Disease Prevention and Health Promotion (Brewer), CDC, Atlanta, Georgia; Los Angeles County Department of Health Services (Fielding); University of Minnesota School of Public Health (Toomey), Minneapolis, Minnesota; and the Schools of Medicine and Public Health (Naimi), Boston University, Boston, Massachusetts

Author affiliations are shown at the time the research was conducted.

Names and affiliations of the Task Force members can be found at www.thecommmunityguide.org/about/task-force-members.html.

Address correspondence to: Robert A. Hahn, PhD, MPH, Community Guide Branch, Epidemiology and Analysis Program Office, CDC, 1600 Clifton Road, Mailstop E-69, Atlanta GA 30333. E-mail: rhahn@cdc.gov.

0749-3797/$36.00

doi: 10.1016/j.amepre.2012.01.002

the estimated economic cost of excessive drinking in the U.S. was $223.5 billion.[5] The reduction of excessive alcohol consumption is thus a matter of major public health and economic interest.

Following the end of Prohibition in the U.S. in 1933, some states continued prohibition at the state level. "Control" states allowed alcohol to be sold, but only through government-run retail stores; "license" states allowed retail sales by commercial interests.[6] In all states, wholesale of alcoholic beverages is under state control. Currently, no states prohibit alcohol sales, and the number of states that have retained control over retail sales has been reduced through privatization. However, in all states with government control over certain beverage types, government control is restricted to *off-premises* sales outlets (i.e., outlets where alcohol is sold for consumption elsewhere); government control does not affect the *on-premises* sale of alcohol in any state.

In the U.S. and Canada, privatization most often affects wine and spirits (e.g., vodka and whiskey). In the U.S., all states and counties that permit the sale of alcohol allow privatized retail sales of beer, and most allow privatized retail sale of all alcoholic beverages. In contrast, in the Scandinavian countries where most other studies of privatization have been conducted, privatization and re-monopolization generally affect beer sales. The National Alcohol Beverage Control Association (NABCA; www.nabca.org) classifies state retail alcohol sales control policies as shown in Table 1.

The predominant trend in the U.S. and elsewhere is toward relaxing government control over the sale of alcoholic beverages, including by privatizing alcohol sales. The formation of the European Union has also led to a loosening of national control and increased privatization in some member nations.[7,8] However, because privatization could plausibly lead to increases in excessive alcohol consumption and related harms, a public health intervention of possible interest to some jurisdictions and decision makers may be the reversal of privatization (*re-monopolization*) or the maintenance of government control where it exists currently.

This review addresses three research questions related to the effect of privatizing retail sales of alcoholic beverages. (1) Does retail privatization of a specific type of alcoholic beverage increase its excessive consumption and associated harms (e.g., alcohol-impaired driving, assaults, and cirrhosis of the liver)? (2) Does privatization of sales of one type of alcoholic beverage also reduce excessive consumption of alcoholic beverages for which sales are not privatized (e.g., does the retail privatization of wine sales lead consumers to reduce their consumption of liquor, if liquor sales are still subject to government control)? (3) Does the re-establishment of state control over the retail sales of an alcoholic beverage (re-monopolization) reduce the excessive consumption of that beverage and the harms related to excessive consumption?

Table 1. The National Alcohol Beverage Control Association (NABCA; www.nabca.org/) classification of state retail alcohol sales control policies

State	Wine	Spirits
Alabama	Private	Government
Idaho	Private	Government and agents
Iowa	Private	Private
Maine	Private	Agents
Maryland (Montgomery County only)	Government	Government
Michigan	Private	Private
Mississippi	Private	Private
Montana	Private	Agents
New Hampshire	Private	Government
North Carolina	Private	Government
Ohio	Private	Agents
Oregon	Private	Agents
Pennsylvania	Government	Government
Utah	Government and agents	Government and agents
Vermont	Private	Agents
Virginia	Private	Government
Washington	Private and government	Government and agents
West Virginia	Private	Private
Wyoming	Private	Private

Findings, Recommendations, and Directives from Other Reviews and Advisory Groups

The 2010 WHO-sponsored review *Alcohol: No Ordinary Commodity*,[8] a consensus narrative review of a broad array of alcohol interventions, assessed government control of the retail sale of alcohol, and rated it 3+ (the highest rating) for effectiveness in reducing excessive alcohol consumption and related harms and for the extent of research supporting the finding and 2+ on cross-cultural testing, suggesting generalizability across settings. The cost of implementing government control was rated as low. The reviewers concluded that "the evidence is quite strong that off-premise government control systems limit alcohol consumption and alcohol-related problems,

and that elimination of government off-premise monopolies can increase total alcohol consumption."[8] The present analysis adds more recent evidence on effects of privatization and applies a more formal protocol to the evaluation and synthesis of available scientific evidence on this topic. The intervention reviewed here may be helpful in addressing several national health objectives related to substance abuse prevention, as specified in *Healthy People 2020*.[9]

Evidence Acquisition

The *Guide to Community Preventive Services* (*Community Guide*) systematic review process was used to assess whether privatization leads to increases in excessive alcohol consumption and related harms. More details on the *Community Guide* review process are presented elsewhere.[10,11] In brief, this process involves forming a systematic review development team; developing a conceptual approach to organizing, grouping, and selecting interventions; prioritizing these interventions; searching for and retrieving the existing research evidence on the effects of the interventions; assessing the quality of each study; abstracting information from each study that meets qualifying criteria; drawing conclusions about the body of evidence on intervention effectiveness; and translating the evidence on effectiveness into recommendations. To help ensure objectivity, the systematic review development team consists of systematic review methodologists and subject matter experts from a range of agencies, organizations, and academic institutions. The review team works under the oversight of the nonfederal, independent Community Preventive Services Task Force (Task Force), which directs the work of the *Community Guide*.

The systematic review development team collects and summarizes evidence on (1) the effectiveness of interventions in altering the health-related outcomes of interest; and (2) benefits and harms of the intervention on other health and nonhealth outcomes. When an intervention is shown to be effective, information is also included about (3) the applicability (or generalizability) of the evidence to diverse population segments and settings; (4) the economic impact of the intervention; and (5) barriers to implementation. The systematic review development team then presents the results of this review process to the Task Force, which considers all of the evidence presented and determines whether it is sufficient to warrant a recommendation for practice or policy.[10] The rules of evidence under which the Task Force makes its determination require consideration of several aspects of the body of evidence, including the number of studies of different levels of design suitability and execution, as well as judgments regarding the consistency of the findings, the public health importance of the overall effect size, and the balance of the outcome of interest with other consequences of the intervention.

Conceptual Approach and Analytic Framework

Six intermediate consequences of privatization are hypothesized to affect alcohol consumption (Figure 1). First, privatization generally leads to increases in the density of off-premises alcohol outlets.[12] Second, privatization may increase the availability of alcohol by increasing the days and hours during which it is sold.[13,14] Third, privatization may increase the availability of specific types and brands of alcoholic beverages.

Fourth, privatization may affect the retail price of alcohol because of various factors that may either increase or decrease prices; overall, prices tend to increase with privatization.[13,14] However, the wider range of alcohol products typically available in privatized systems may result in more low-priced products that would appeal to high-volume or high-risk drinkers, even if the average price for all alcohol products were similar in state-owned and privatized systems. Fifth, because it introduces competition among alcohol outlets, privatization may lead to increased alcohol advertising in various venues, including TV, radio, billboards, and at the point of purchase.[13] Sixth, because of increased numbers of outlets and less direct governmental control, privatization may lead to decreased enforcement of and compliance with sales regulations (e.g., minimum-drinking-age laws).[15]

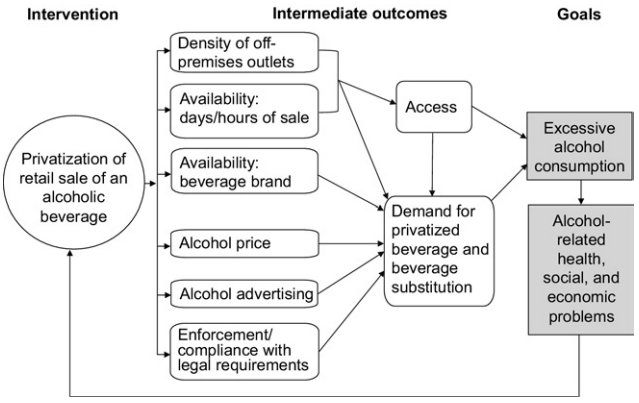


Figure 1. Analytic framework

Note: Privatization of retail sales of alcoholic beverages: Oval indicates intervention; rectangles with rounded corners indicate mediators or intermediate outcomes; rectangles indicate health outcomes.

Many of these consequences of privatization may increase the demand for the privatized beverage, substitution to or from other beverages, and access to alcoholic beverages. In turn, these consequences may affect excessive consumption and related harms. In contrast, re-monopolization is expected to have effects opposite to those of privatization.

Inclusion and Exclusion Criteria

To qualify as a source of primary evidence for this review, a study had to meet several criteria. It had to evaluate retail privatization or renewed government control; be conducted in a country with a high-income economy[16]; be primary research (rather than a review of other research); be published in English; and have a comparison group and/or compare conditions before and after privatization. Optimally, alcohol consumption in locations that experienced changes in government control (e.g., privatization) would be compared with alcohol consumption in similar locations that did not experience this change.

A study also had to report outcomes related to excessive alcohol consumption or related harms. Some specific harms of interest included alcohol-related medical conditions (e.g., liver cirrhosis), alcohol-impaired driving and alcohol-related crashes, unintentional injuries, intentional injuries, and crime. Government reports were included, but unpublished papers were not, as they may be difficult for others to access and are not peer-reviewed.

When studies assessed multiple outcomes of interest, those with the strongest known association with excessive alcohol consumption were given more weight. Outcome measures in the studies that had the strongest association with excessive alcohol consumption included binge drinking, heavy drinking, liver cirrho-

sis mortality, alcohol-related hospital admissions, and alcohol-related motor vehicle crashes (or strong proxies for such crashes, e.g., single-vehicle nighttime crashes for alcohol-related motor vehicle crashes).[17] Less-direct measures included per capita ethanol consumption, a well-recognized proxy for estimating the number of excessive drinkers in a population[8,18,19]; unintentional injuries; suicide; and crime (e.g., homicide and aggravated assault).

Cross-sectional and panel studies in which the sequence of events was not taken into account (e.g., counting year as an ordinal variable) were considered secondary evidence in this review. Secondary evidence was regarded as useful for generating hypotheses and for strengthening or weakening conclusions based on primary evidence, but was insufficient alone for assessing intervention effectiveness. Many of the studies qualifying for this review used per capita alcohol sales as a proxy for excessive drinking. In assessing the quality of study execution, a penalty (described below) was assigned to studies that assessed changes in population-level consumption rather than changes in excessive drinking by individuals. This was a conservative approach because excessive drinking and per capita alcohol consumption are strongly related both theoretically and empirically.[18–20]

This empirical relationship is conceptualized in the "single distribution theory," which asserts that excessive drinkers, including binge and heavy drinkers, account for a consistent proportion of the drinking population in a given setting,[19] such that the prevalence of excessive alcohol consumption is directly related to per capita alcohol sales. Further, the theory proposes that the relationship between per capita sales and "heavy" or excessive drinking is quadratic, meaning that, "If population A has twice the average consumption of population B, then A has about four times (i.e., 2^2) the prevalence of heavy drinking." Cook and Skog[19] report evidence from multiple countries supporting this proposition. Thus, changes in per capita alcohol consumption would be expected to have a greater effect on excessive alcohol consumption, including binge drinking, than on non-excessive alcohol consumption. This evidence supports the Task Force's use of per capita alcohol consumption based on sales as an outcome measure for assessing the impact of privatization on excessive alcohol consumption.

Search for Evidence

The following databases were searched for this review: Econlit, PsycINFO, Sociology Abstracts, MEDLINE, Embase, and EtOH. All years were searched up to October 2007. (Details of the search strategy are available at www.thecommunityguide.org/alcohol/supportingmaterials/SSalcoholuse.html.) This search was updated in the ISI Web of Knowledge through December 2010 by the Alcohol Epidemiology Program at the University of Minnesota. Reference lists of articles reviewed as well as lists in review articles were searched, and subject matter experts were consulted.

Assessing the Quality and Summarizing the Body of Evidence on Effectiveness

Each study that met the criteria for candidate studies was read by two reviewers, who used standardized criteria (available at www.thecommunityguide.org/about/methods.html) to assess suitability of study design and threats to validity. Uncertainties and disagreements between reviewers were reconciled by consensus among the team members. The team's classification of the designs of studies reviewed corresponds with the research questions of the review and the standards of the *Community Guide* review process[11] and may differ from the classification reported in the original studies.

The quality of studies that were candidates as primary evidence for this review was evaluated both in terms of design and execution. Studies with greatest design suitability were those in which data on exposed and comparison populations were collected prospectively; studies with moderate design suitability were those in which data on exposed and comparison populations were collected retrospectively or in which there were multiple pre- or post-intervention measurements, but no concurrent comparison population; and studies with least-suitable designs were cross-sectional studies or those in which there was no comparison population or only a single pre- and post-measurement in the intervention population. On the basis of the number of threats to validity—such as poor measurement of exposure or outcome, lack of control of potential confounders, or high attrition—studies were characterized as having good (at most one threat to validity), fair (two to four threats), or limited (five or more threats) quality of execution. Studies with good or fair quality of execution and any level of design suitability (greatest, moderate, or least) qualified for the body of evidence.

Effect estimates were calculated as relative percentage change using the following formulas:

- For studies with before–after measurements and concurrent comparison groups:

$$\text{Effect estimate} = \left[\left(I_{post}/C_{post}\right)/\left(I_{pre}/C_{pre}\right) - 1\right] \times 100\,\%$$

where:

I_{post} = last reported outcome rate or count in the intervention group after the intervention;
I_{pre} = last reported outcome rate in the intervention group before the intervention;
C_{post} = last reported outcome rate in the comparison group after the intervention;
C_{pre} = last reported outcome rate in the comparison group before the intervention.

- For studies with before–after measurements but no concurrent comparison:

$$\text{Effect estimate} = \left[\left(I_{post} - I_{pre}\right)/I_{pre}\right] \times 100\,\%.$$

Several events of privatization (e.g., the privatization of wine in Iowa in 1985) were assessed by more than one team of researchers, thus resulting in multiple studies of the same event. Effect estimates are reported for each research group, noting which were associated with a single event. Median effect sizes are calculated using the means of privatization events with differing findings from different researchers.

Evidence Synthesis

Intervention Effectiveness

The effects of 12 distinct privatization events were assessed in 17 studies and reported in 13 publications.[14,21–32] In addition, there was one study of re-monopolization, described separately below.[33] The privatization events assessed were in seven U.S. states (Alabama, Idaho, Iowa [two events], Maine, Montana,

New Hampshire, West Virginia); two Canadian provinces (Quebec [two events] and Alberta); and Finland. Several publications described a single privatization event, and several publications each assessed more than one privatization event.

All studies used alcohol sales data as an index of population-level alcohol consumption except one[25] that assessed changes in individual-level consumption (in Finland). Fitzgerald and Mulford[14,22] also assessed changes in self-reported consumption in addition to sales data (in Iowa). However, as they note, their measures of alcohol consumption were problematic (e.g., separate cross-sectional studies collected under different sampling procedures with different interview procedures), and thus only the assessments of changes in alcohol sales from their study are included in this review.

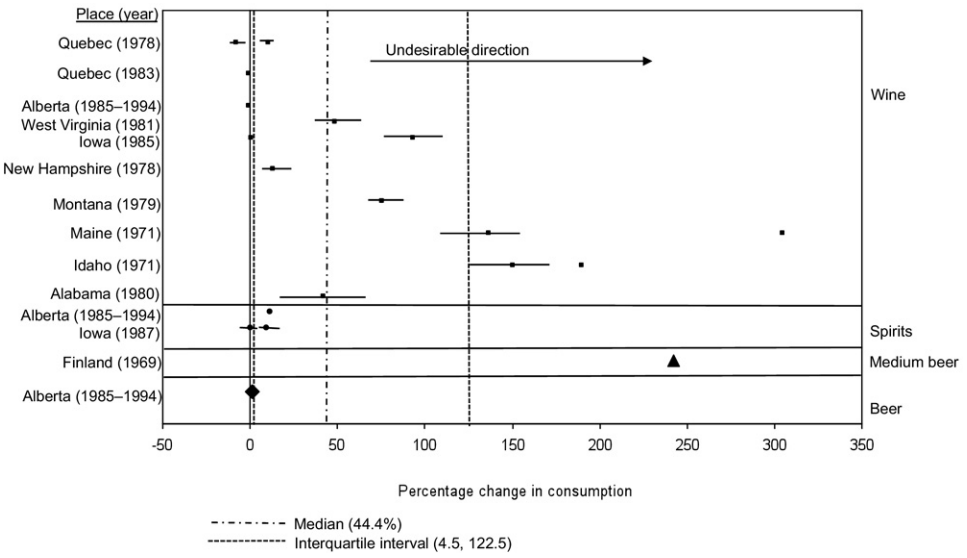


Figure 2. Percentage change by location of privatization event in consumption of privatized alcoholic beverages attributable to privatization

Note: Bars on the point estimate represent 95% CIs when reported or calculated.

In the U.S., privatized beverages were limited to wine and spirits, as beer was already privatized. In Canada, sale of beer was privatized in addition to sales of wine and spirits, and in Scandinavia, privatization (and re-monopolization) focused on beer sales. The privatization events assessed in these studies occurred between 1950 and 2000. Three studies used autoregressive integrated moving average (ARIMA) time series.[27–29] Fourteen studies (presented in eight publications)[14,23,26–30,32] were of greatest design suitability; three studies (presented in two publications)[24,25] were of moderate design suitability. All studies were of fair quality of execution.

The qualifying studies provided information on several of the intermediate consequences of privatization discussed above (Figure 1). These consequences include increased numbers of alcoholic beverage outlets, increased hours and days of sale, advertising, greater brand selection, and acceptance of alternate forms of payment (e.g., credit card).[13,14,23,26–30,32,34] All of these intermediate outcomes would be expected to result in increased consumption.

Most studies reviewed reported generally higher prices for alcoholic beverages in the privatized than in the state control setting. The higher prices may be the consequence of relative inefficiencies of scale (e.g., multiple smaller outlets and increased overhead expenses).[35] In contrast to other intermediate outcomes, higher prices would be expected to result in decreased consumption. However, Fitzgerald and Mulford[21] assessed whether the increase in the price of spirits following privatization of retail sales in Iowa had affected consumers' purchasing behavior, and found that only 37.4% of those surveyed who purchased liquor in the past month recognized that prices had increased, and <2% of Iowa consumers reported increasing their purchase of liquor from adjacent states.

Effects of privatization on consumption of privatized beverages. Overall, the median increase in per capita sales of privatized beverages was 44.4%, with an interquartile interval of 4.5% to 122.5% (Figure 2). Some differences, however, were observed across studies in the impact of privatization on retail sales of privatized beverages and in the relationship between privatization and other public health outcomes, described below.

Studies assessing the effects of the privatization of wine (1985) and spirits (1987) sales in Iowa had inconsistent findings. Wagenaar and Holder[31] reported that wine consumption increased 93.0% (95% CI=69.3%, 120.2%) from baseline to 44 months after privatization of wine sales in Iowa, with no decrease in spirits or beer consumption. Following the subsequent privatization of spirits sales in Iowa 2 years later, these researchers[23] reported a 9.5% (95% CI=3.5%, 15.9%) increase in spirits consumption, along with a 12.1% (95% CI= −20.6%, −2.7%) decrease in wine consumption and no change in

beer consumption. They[23,31] also found no evidence that privatization affected alcohol purchasing across state lines (effect estimate 0.1%, 95% CI= −3.1%, 6.2%).

In contrast, Mulford and Fitzgerald[27] found that wine privatization was associated with a nonsignificant long-term increase of only 0.5% (95% CI= −6.8%, 8.3%) in wine sales, and that spirits privatization was associated with a nonsignificant long-term increase of 0.7% (95% CI= −1.9%, 3.4%) in spirits sales. For both beverages, the nonsignificant long-term effects were preceded by 3-year spikes in sales. The differences between the conclusions of these two research teams about the effects of the privatization of wine and spirits in Iowa may be the consequence of different modeling strategies, different time periods covered, and different forms of alcoholic beverages included—particularly the inclusion of "wine coolers" in measures of wine sales by Wagenaar and Holder.[31,32]

Finally, Makela[25] assessed the impact on alcohol consumption of a law in Finland that allowed the sale of medium-strength beer (MSB) in grocery stores. This was the only study included in the review that assessed changes in self-reported alcohol consumption by individuals over time. Survey participants were specifically asked about their levels of consumption before implementation of the new law and then again in the year following its implementation. The researchers stratified their findings based on drinking patterns of respondents before and after privatization. Consumption of all alcoholic beverages (not just the privatized beverage) increased by a mean of 1.7 L of pure alcohol per year per person interviewed (approximately 137 ounces of 80° proof liquor). Makela reports that 86% of the increase in overall alcohol consumption was attributable to increases in the privatized beverage (MSB). The greatest increase in alcohol consumption after privatization was observed among those who reported drinking between 17 and 68 ounces of pure alcohol per year at first interview. However, there was also an increase in consumption in the population that reported no alcohol consumption within the past 30 days when first interviewed.

Effects of privatization on alcohol-related harms. Two studies assessed the association between retail privatization and motor vehicle crashes. One study estimated that incremental privatization over a 20-year period was associated with a nonsignificant 11.3% (95% CI= −33.9%, 19.0%) decrease in traffic fatalities in Alberta, Canada.[29] This study estimated the degree of privatization over a long period preceding final privatization in 1994, had only 1 year of follow-up, and used a proxy outcome measure.

A second study assessed changes in alcohol-related harms associated with the Iowa privatization of wine in 1985 and spirits in 1987.[14,26] The researchers compared the period before 1985 with the period after 1989, when both wine and spirits sales were privatized. Despite increased per capita sales of both wine and spirits, there was a reported 1.6% decline in nighttime motor vehicle crashes and a 5.5% decline in liver cirrhosis. However, initial mortality data were for 1985—the same year in which the privatization of wine occurred; thus these data included deaths both prior to and following privatization, weakening the analysis. Moreover, no comparison data were provided to adjust for national or regional trends in these outcomes over the time period evaluated.

Effects of privatization on the consumption of *nonprivatized* alcoholic beverages. Many of the studies[14,25,26,31,32,36] reviewed also assessed the effect of privatizing the sale of one type of alcoholic beverage on the sale of other nonprivatized beverages. In the seven settings assessed, the sales of nonprivatized alcoholic beverages decreased a median of 2.2%, with an interquartile interval ranging from a decrease of 6.6% in sales to a decrease of 0.1% (Figure 3). These decreases are not of sufficient magnitude to offset the overall increase in per capita sales of privatized beverages.

Effects of re-monopolization on alcohol-related outcomes. One study in Sweden[33] directly assessed effects of a 1977 re-monopolization of the sale of MSB (2.26%–3.50% alcohol by volume; beer in the U.S. is generally 4%–6% alcohol by volume). The study was of moderate design suitability and fair execution. Re-monopolization resulted in a substantial decline in the number of outlets for MSB, from 11,550 to 300. The effects of this policy change on hospitalization for several alcohol-related outcomes (alcoholism, alcohol intoxication, alcoholic psychosis, hospitalizations for acute alcohol intoxication, suicides, falls, motor vehicle crashes, and assaults) were assessed using time-series design, comparing the 4 years before and after re-monopolization. The results were stratified by four age categories (10–19 years, 20–39 years, 40–59 years, and ≥60 years), and the study did not provide data to allow aggregation across age groups.

The researchers identified a number of positive changes in health outcomes following re-monopolization. Hospital admissions for the treatment of alcoholism, alcohol intoxication, and alcohol psychosis decreased across all age groups ($p>0.05$), and there was a 20% decline ($p<0.05$) in these outcomes among people aged 10–19 years. Hospitalizations for acute alcohol intoxication decreased across all ages from 3.5% to 14.7% ($p>0.05$). Suicides decreased from 1.7% to 11.8% ($p>0.05$). Falls decreased from 3.6% to 4.9% ($p>0.05$). Motor vehicle crashes decreased 14% ($p<0.05$) for three age categories (10–19 years, 40–59 years, and ≥60 years) and by 4.4% for those aged 20–39 years ($p>0.05$). In contrast, assaults *increased* from 6.9% to 14.8% ($p>0.05$)

in three of four age categories, and decreased by 1.4% among those aged 20–39 years ($p>0.05$). In summary, the re-monopolization of MSB in Sweden was associated with reductions in most of the alcohol-related harms assessed across all age groups; however, many of these effects were not significant.

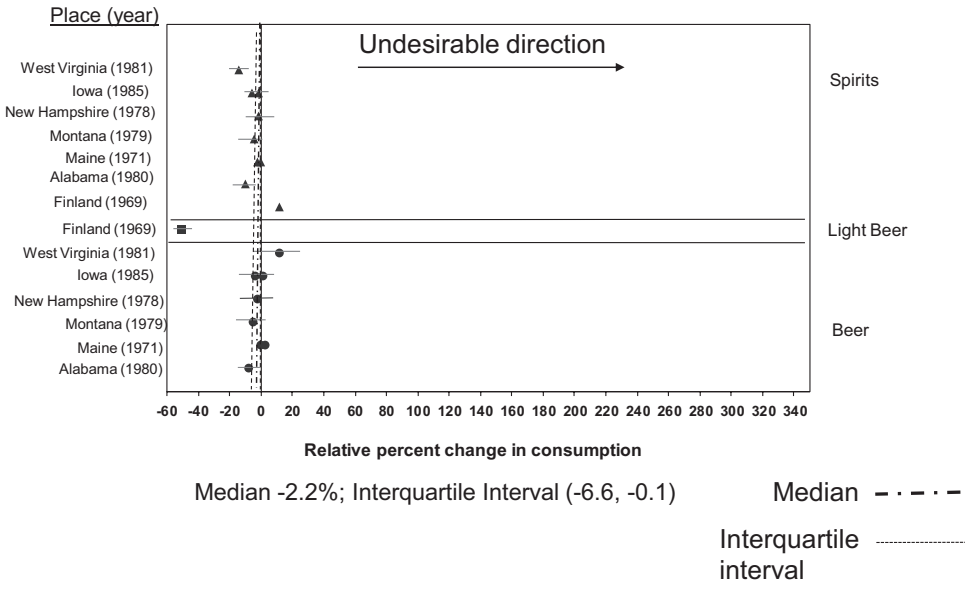


Figure 3. Percentage change in consumption of nonprivatized alcoholic beverages attributable to privatization of other alcoholic beverages

Note: Bars on the point estimate represent 95% CIs when reported or calculated.

Cross-sectional studies, panel studies, and summarized studies in translation. The team found 20 cross-sectional and panel studies[29,35,37–54] assessing the association of privatization with alcohol consumption and related harms; five studies[35,38,43,47,48] had multiple outcomes. (Panel studies were included in this section if they did not include time as an ordinal covariate or compare alcohol-related outcomes before and after events of privatization.) Seventeen studies[29,35,37,38,40,42–48,50–54] found that privatization was associated with greater consumption (nine were significant,[37,40,42–44,46,48,52,53] seven were not,[29,35,38,45,47,50,54] and one[51] did not report significance). Four studies found decreased consumption (three were significant,[40,43,44] one not[38]). Three studies[37,41,50] assessed the association of privatization with cirrhosis mortality; all were positive and two[37,41] were significant. Finally, two studies assessed the association of privatization with motor vehicle fatalities,[39,46] and both found a positive, but nonsignificant association; one study found a negative, significant association of privatization and drunk driving. Overall, this evidence is consistent with evidence from primary studies indicating positive associations between privatization and increased population-level consumption and between privatization and alcohol-related harms.

Makela, Rossow, and Tryggvesson[36] published (in English) a review of studies conducted in Finland, Sweden, and Norway that were not published in English translation and thus were not included in this review. These studies examined the effect of either privatizing or re-monopolization of the sales of medium or strong beer between the mid-1960s and the early 1990s on various kinds of alcohol-related harms (e.g., arrests for drunkenness and alcohol-related illnesses).

The researchers found that when beer sales were privatized, there were increases in alcohol consumption and alcohol-related harms such as arrests for drunkenness and alcohol-related illnesses. They also found that re-monopolization generally resulted in decreased population-level consumption of the affected alcoholic beverage, decreases in excessive alcohol consumption, and decreases in alcohol-related harms. In addition, they found that when a particular strength of beer became more or less accessible, consumers tended to purchase the beverage type that was more readily accessible. This beverage substitution effect appeared to be stronger among different strengths of beer than among different types of alcoholic beverages.

Summary of Intervention Effectiveness

Across the studies qualifying for this review, the privatization of off-premises retail sales of an alcoholic beverage was associated with a median 44.4% increase (interquartile interval 4.5%, 122.5%) in the per capita sales of the beverage and with a 2.2% decline (interquartile interval −6.6%, −0.1%) in the per capita sales of beverages for which sales were not privatized. One study (in Finland)[25] found that the increases in consumption occurred among drinkers at all consumption levels. A single study[33] evaluating the effects of re-monopolization of alcohol sales found that this change was associated with a subsequent decrease in several alcohol-related harms (e.g., hospitalizations for acute alcohol intoxication).

Other Harms and Benefits

Government control over retail alcohol sales generally results in lower alcohol outlet density. In addition to potential public health benefits, lower outlet density may improve quality of life by reducing property damage and public disturbance (e.g., public intoxication). However, the studies reviewed did not assess these effects. The review team did not postulate any serious harms associated with the maintenance of government control over retail alcohol sales.

Applicability

Consistent evidence of the association between privatization and increased per capita alcohol sales comes from studies done in multiple settings in the U.S., Canada, and Europe. Most of these studies evaluated the effects of privatizing the sales of wine and spirits. Only one Swedish study[33] specifically assessed the impact of re-monopolization (of MSB). The findings from the current review apply most directly to the impact of privatizing the sale of wine and spirits in high-income nations such as the U.S.

Economic Efficiency

The present systematic economic review identified one study[55] in Canada that used simulation modeling to estimate healthcare and law enforcement costs and costs of lost productivity due to disability and premature mortality in the event all Canadian provinces and territories were to privatize alcohol sales. The study was judged to be satisfactory by *Community Guide* economics criteria (www.thecommunityguide.org/about/EconAbstraction_v5.pdf). Study authors concluded that these costs were substantially greater than the tax and mark-up revenue gained from increased sales associated with privatization; however, benefit data were not documented.

Research Gaps

Although the studies reviewed have demonstrated an association between privatization and increases in the per capita consumption of the privatized beverages without substantial reductions in consumption of other alcoholic beverages, additional research is needed to clarify the relationship between privatization and various patterns of excessive alcohol consumption (e.g., binge drinking) as well as harms related to it. Most useful would be cohort studies in the U.S. similar to the one conducted by Makela et al.[25] in Finland, assessing the effects of privatization on patterns of excessive alcohol consumption (e.g., binge drinking) and alcohol-related harms. It would also be useful to evaluate the impact of increased government control over alcohol sales (e.g., re-monopolization) on excessive alcohol consumption and related harms, were such events to occur in the U.S. or other high-income nations.

Privatization has assumed different forms in different states and localities. Thus, it would be useful to determine how the effects of privatization observed in this review vary by the degree of government regulation and other specific parameters of the privatization. Although, in general, government control establishes a greater degree of regulation over retail alcohol sales than systems in which sales have been privatized, Her et al.[13] have noted that "privatization might involve a change from a very restrictive alcohol management system to a loosely regulated private one; it would also potentially involve a change from a commercially orientated public system to a private sector operation that is heavily regulated."

No peer-reviewed studies were found that evaluated economic effects of privatizing the sale of alcoholic beverages in the U.S. The anticipated effects of privatization include a large, but short-term, source of revenue to states; a potential increase in healthcare and criminal justice costs; and productivity losses from expected increases in excessive alcohol consumption owing to greater availability and/or lower prices. Studies assessing these economic impacts would be useful for informing future discussions of this issue. It would be useful to assess the effects of different specific approaches to privatization on state revenues associated with sales and taxes on alcoholic beverages.

Conclusion

The evidence consistently showed that privatization of retail alcohol sales was associated with a substantial increase in per capita sales of the privatized beverages, a well-established proxy for excessive alcohol consumption. There was also evidence that re-monopolization is associated with a decrease in alcohol-related harms. Therefore, according to *Community Guide* rules of evidence, there was strong evidence that retail privatization of alcohol sales leads to increases in excessive alcohol consumption.

In the U.S., many states have privatized the retail sales of alcoholic beverages. Currently, three states control the off-premises consumption retail sales of both wine and spirits, and an additional ten states maintain control over the retail sale of spirits alone. In addition, one county in the state of Maryland has county-level control over the retail sale of spirits and wine. The findings of the present report are based solely on evidence related to the public health consequences of privatization, which may be one of several factors considered in making decisions on whether to privatize retail alcohol sales. The maintenance of government control of off-premises sale of alcoholic

beverages is one of many effective strategies to prevent or reduce excessive consumption, which is one of the leading causes of preventable death and disability.

We would like to thank Steve Wing (SAMHSA) and Ralph Hingson, ScD, MPH, National Institute of Alcohol Abuse and Alcoholism, for participation in this review. We also thank our team of consultants for their comments: Edward Bernstein, MD, Boston University School of Public Health; Richard J. Bonnie, LLB, University of Virginia; Raul Caetano, MD, MPH, PhD, University of Texas at Dallas; Frank Chaloupka, PhD, University of Illinois at Chicago; Norm Giesbrecht, PhD, Centre for Addiction and Mental Health, Canada; Thomas Greenfield, PhD, National Alcohol Research Center/Alcohol Research Group, California; Joel Grube, PhD, and Harold Holder, PhD, Pacific Institute for Research and Evaluation (PIRE), California; Mel Kohn, MD, Oregon Department of Human Services; Jurgen Rehm, PhD, Centre for Addiction and Mental Health, Canada; Robin Room, PhD, Stockholm University; Henry Saffer, PhD, National Bureau of Economic Research; Leslie Snyder, PhD, University of Connecticut; Tim Stockwell, PhD, University of Victoria; Alex Wagenaar, PhD, University of Florida College of Medicine.

No financial disclosures were reported by the authors of this paper.

References

1. CDC. Alcohol-attributable deaths and years of potential life lost—U.S., 2001. MMWR Morb Mortal Wkly Rep 2004;53(37):866–70.
2. Miller JW, Naimi TS, Brewer RD, Jones SE. Binge drinking and associated health risk behaviors among high school students. Pediatrics 2007;119(1):76–85.
3. National Center for Health Statistics. Health, U.S., 2010, with special feature on death and dying. www.cdc.gov/nchs/data/hus/hus10.pdf.
4. Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE. Monitoring the future: National survey results on drug use, 1975–2008. Volume II, College students and adults ages 19–50. Bethesda MD: National Institute on Drug Abuse, 11-1-2009. NIH Publication No. 09-7403.
5. Bouchery EE, Harwood HJ, Sacks JJ, Simon CJ, Brewer RD. Economic costs of excessive alcohol consumption in the U.S., 2006. Am J Prev Med 2011;41(5):516–24.
6. Aaron P, Musto D. Temperance and prohibility in America: an historical overview. In: Moore MH, Gerstein DR, editors. Alcohol and public policy: beyond the shadow of prohibition. Washington DC: National Academy Press, 1981:127–81.
7. Holder HD, Giesbrecht N, Horverak O, et al. Potential consequences from possible changes to Nordic retail alcohol monopolies resulting from European Union membership. Addiction 1995;90(12):1603–18.
8. Babor TF. Alcohol: no ordinary commodity. Research and public policy. 2nd ed. New York: Oxford University Press, 2010.
9. DHHS. Healthy People 2020. www.healthypeople.gov/2020/default.aspx.
10. Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based Guide to Community Preventive Service—methods. Am J Prev Med 2000;18(1S):35–43.
11. Zaza S, Wright-De Aguero LK, Briss PA, et al. Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. Am J Prev Med 2000;18(1S):44–74.
12. Campbell AC, Hahn RA, Elder RA, et al.; Task Force on Community Preventive Services. The effectiveness of limiting alcohol outlet density as a means of reducing excessive alcohol consumption and alcohol-related harms. Am J Prev Med 2009;37(6):556–69.
13. Her M, Giesbrecht N, Room R, Rehm J. Privatizing alcohol sales and alcohol consumption: evidence and implications. Addiction 1999;94(8):1125–39.
14. Fitzgerald JL, Mulford HA. Consequences of increasing alcohol availability: the Iowa experience revisited. Br J Addict 1992;87(2):267–74.
15. Rossow I, Karlsson K, Raitasalo K. Old enough for a beer? Compliance with minimum legal age for alcohol purchases in monopoly and other off-premise outlets in Finland and Norway. Addiction 2008;103:1468–73.
16. World Bank. World Development Indicators 2006. devdata.worldbank.org/wdi2006/contents/cover.htm.
17. Gruenewald PJ, Millar AB, Treno AJ, Yang Z, Ponicki WR, Roeper P. The geography of availability and driving after drinking. Addiction 1996;91(7):967–83.
18. Nelson D, Naimi TS, Brewer RD, Roeber J. U.S. State alcohol sales compared to survey data, 1993–2006. Addiction 2010;105:1589–96.
19. Cook PJ, Skog OJ. Alcool, alcoolisme, alcoolisation— comment. Alcohol Health Res World 1995;19(1):30–1.
20. Greenfield TK, Rogers JD. Who drinks most of the alcohol in the U.S.? The policy implications. J Stud Alcohol 1999;60(1):78–89.
21. Fitzgerald JL, Mulford HA. Privatization, price and cross-border liquor purchases. J Stud Alcohol 1993;54(4):462–4.
22. Fitzgerald JL, Mulford HA. Alcohol availability, drinking contexts and drinking problems: the Iowa experience. J Stud Alcohol 1993;54(3):320–5.
23. Holder HD, Wagenaar AC. Effects of the elimination of a state monopoly on distilled spirits' retail sales: a time-series analysis of Iowa. Br J Addict 1990;85(12):1615–25.
24. Macdonald S. The impact of increased availability of wine in grocery stores on consumption: four case histories. Br J Addict 1986;81:381–7.
25. Makela P. Whose drinking does the liberalization of alcohol policy increase? Change in alcohol consumption by the initial level in the Finnish panel survey in 1968 and 1969. Addiction 2002;97(6):701–6.
26. Mulford HA, Fitzgerald JL. Consequences of increasing off-premise wine outlets in Iowa. Br J Addict 1988;83(11):1271–9.
27. Mulford HA, Ledolter J, Fitzgerald JL. Alcohol availability and consumption: Iowa sales data revisited. J Stud Alcohol 1992;53(5):487–94.
28. Smart RG. The impact on consumption of selling wine in grocery stores. Alcohol Alcohol 1986;21(3):233–6.
29. Trolldal B. An investigation of the effect of privatization of retail sales of alcohol on consumption and traffic accidents in Alberta, Canada. Addiction 2005;100(5):662–71.
30. Trolldal B. The privatization of wine sales in Quebec in 1978 and 1983 to 1984. Alcohol Clin Exp Res 2005;29(3):410–6.
31. Wagenaar AC, Holder HD. A change from public to private sale of wine: results from natural experiments in Iowa and West Virginia. J Stud Alcohol 1991;52(2):162–73.
32. Wagenaar AC, Holder HD. Changes in alcohol consumption resulting from the elimination of retail wine monopolies: results from five U.S. states. J Stud Alcohol 1995;56(5):566–72.
33. Ramstedt M. The repeal of medium-strength beer in grocery stores in Sweden: the impact on alcohol-related hospitalizations in different age groups. In: Room R, ed. The effects of Nordic alcohol policies: what happens to drinking and harm when alcohol controls change? Finland: Nordic Council for Alcohol and Drug Research (NAD), 2002:69–78.
34. Mustonen H, Sund R. Changes in the characteristics of drinking occasions resulting from liberalization of alcohol availability: A reanalysis of the 1968 and 1969 Finnish panel survey data. In: Room R, ed. The effects of Nordic alcohol policies: what happens to drinking and harm when alcohol controls change? Nordic Council for Alcohol and Drug Research, 2002:49–55.

35. Simon JL. The economic effects of state monopoly of packaged-liquor retailing. J Polit Econ 1966;74:188–94.
36. Makela P, Rossow I, Tryggvesson K. Who drinks more and less when policies change? The evidence from 50 years of Nordic studies. In: Room R, editor. The effects of Nordic alcohol policies: what happens to drinking and harm when alcohol controls change? Finland: Nordic Council for Alcohol and Drug Research, 2002:42–55.
37. Andreano RL, Li JY. Death rates from liver cirrhosis in alcohol monopoly and free-market states. Q J Stud Alcohol 1974;35(Pt A):649–54.
38. Beard TR, Gant PA, Saba RP. Border-crossing sales, tax avoidance, and state tax policies: an application to alcohol. South Econ J 1997; 64(1):293–306.
39. Colon I. The influence of state monopoly of alcohol distribution and the frequency of package stores on single motor vehicle fatalities. Am J Drug Alcohol Abuse 1982;9(3):325–31.
40. Goel RK, Morey MJ. The interdependence of cigarette and liquor demand. South Econ J 1995;62(2):451–9.
41. Heien D, Pompelli G. Stress, ethnic and distribution factors in a dichotomous response model of alcohol abuse. J Stud Alcohol 1987; 48(5):450–5.
42. Hoadley JF, Fuchs BC, Holder HD. The effect of alcohol beverage restrictions on consumption: a 25-year longitudinal analysis. Am J Drug Alcohol Abuse 1984;10(3):375–401.
43. Kenkel DS. Drinking, driving, and deterrence: The effectiveness and social costs of alternative policies. J Law Econ 1993;36(2):877–913.
44. Kenkel DS. New estimates of the optimal tax on alcohol. Economic Inquiry 1996;34(2):296–319.
45. McCornac DC, Filante RW. The demand for distilled spirits: an empirical investigation. J Stud Alcohol 1984;45(2):176–8.
46. Miller T, Snowden C, Birckmayer J, Hendrie D. Retail alcohol monopolies, underage drinking, and youth impaired driving deaths. Accid Anal Prev 2006;38(6):1162–7.
47. Nelson JP. State monopolies and alcoholic beverage consumption. J Regulatory Econ 1990;2(1):83–98.
48. Nelson JP. Advertising bans, monopoly, and alcohol demand: testing for substitution effects using state panel data. Rev Industrial Org 2003;22(1):1–25.
49. Ornstein S, Hanssens D. Alcohol control laws and the consumption of distilled spirits and beer. J Consum Res 1985;12(2):200–13.
50. Popham RE, Schmidt W, de Lint J. Government control measures to prevent hazardous drinking. In: Ewing JA, Rouse BA, editors. Drinking: alcohol in American society: issues and current research. Chicago: Nelson-Hall, 1978:239–66.
51. Spellman WE, Jorgenson MR. Liquor control and consumption. J Stud Alcohol 1983;44(1):194–7.
52. Stockwell T, Zhang J, Macdonald S, Pakula B, Gruenewald P, Holder H. Changes in per capita alcohol sales during the partial privatization of British Columbia's retail alcohol monopoly 2003–2008: a multi-level local area analysis. Addiction 2009;104:1827–36.
53. Trolldal B. Alcohol price elasticities in control and license states in the U.S., 1982-99. Addiction 2005;100(8):1158–65.
54. Zardkoohi A, Sheer A. Public versus private liquor retailing: an investigation into the behavior of the state governments. South Econ J 1984;50(4):1058–76.
55. Popova S, Patra J, Sarnocinska-Hart A, Gnam WH, Giesbrecht N, Rehm J. Cost of privatization versus government alcohol retailing systems: Canadian example. Drug Alcohol Rev 2011;Epub ahead of print.

Massachusetts Reform and Disparities in Inpatient Care Utilization

Amresh D. Hanchate, PhD,† Karen E. Lasser, MD, MPH,†‡ Alok Kapoor, MD, MSc,† Jennifer Rosen, MD,§ Danny McCormick, MD, MPH,‖ Meredith M. D'Amore, MPH,† and Nancy R. Kressin, PhD*†*

Background: The 2006 Massachusetts health reform substantially decreased uninsurance rates. Yet, little is known about the reform's impact on actual health care utilization among poor and minority populations, particularly for receipt of inpatient surgical procedures that are commonly initiated by outpatient physician referral.

Methods: Using discharge data on Massachusetts hospitalizations for 21 months before and after health reform implementation (7/1/2006–12/31/2007), we identified all nonobstetrical major therapeutic procedures for patients aged 40 or older and for which $\geq 70\%$ of hospitalizations were initiated by outpatient physician referral. Stratifying by race/ethnicity and patient residential zip code median (area) income, we estimated prereform and postreform procedure rates, and their changes, for those aged 40–64 (nonelderly), adjusting for secular changes unrelated to reform by comparing to corresponding procedure rate changes for those aged 70 years and above (elderly), whose coverage (Medicare) was not affected by reform.

Results: Overall increases in procedure rates (among 17 procedures identified) between prereform and postreform periods were higher for nonelderly low area income (8%, $P = 0.04$) and medium area income (8%, $P < 0.001$) cohorts than for the high area income cohort (4%); and for Hispanics and blacks (23% and 21%, respectively; $P < 0.001$) than for whites (7%). Adjusting for secular changes unrelated to reform, postreform increases in procedure utilization among nonelderly were: by area income, low = 13% (95% confidence interval (CI) = [9%, 17%]), medium = 15% (95% CI [6%, 24%]), and high = 2% (95% CI [−3%, 8%]); and by race/ethnicity, Hispanics = 22% (95% CI [5%, 38%]), blacks = 5% (95% CI [−20%, 30%]), and whites = 7% (95% CI [5%, 10%]).

Conclusions: Postreform use of major inpatient procedures increased more among nonelderly lower and medium area income populations, Hispanics, and whites, suggesting potential improvements in access to outpatient care for these vulnerable subpopulations.

Key Words: health reform, disparities, utilization, inpatient care, access to care, socioeconomic status, race, ethnicity

(*Med Care* 2012;50: 569–577)

A central policy assumption in the United States is that expanding health insurance coverage will improve access to health care and outcomes, and make each more equitable for all Americans.[1] Massachusetts (MA) is the site of a key policy-relevant natural experiment[2–8]; recent legislation has resulted in nearly all (96.5%) of the state's residents obtaining health insurance.[9] However, little is known about MA reform's impact on health care utilization, particularly among poor and minority populations, whose access to care the reform sought to increase.

The number of uninsured MA residents fell sharply after the reform was implemented.[10] Among adults aged 18–64, the population targeted by the reform, uninsurance rates declined from 8.4% (2006) to 3.4% (2009) overall, but from 18%–9% among the poorest population quintile, from 15%–5% among blacks, and 20%–13% among Hispanics.[9,11,12] However, the limited evidence of the reform's impact on access to and use of health care, based largely on population surveys, provides a mixed picture. Self-reported rates of a usual source of care and of preventive care visits

From the *VA Boston Healthcare System; †Section of General Internal Medicine, Boston University School of Medicine; ‡Department of Community Health Sciences, Boston University School of Public Health; §Department of Surgery, Boston University School of Medicine, Boston; and ‖Harvard Medical School, Department of Medicine, Cambridge Health Alliance, Cambridge, MA.

The authors declare no conflict of interest.

Supported by NIH grants (1R21NS062677, A.D.H., PI and 1U01HL105342-01, N.R.K., PI). N.R.K. is supported in part by a Research Career Scientist award from the Department of Veterans Affairs, Health Services Research and Development Service (RCS 02-066-1). A.D.H. had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Presented at: Plenary session—oral presentation at the Society of General Internal Medicine, Minneapolis, MN, April 29, 2010; poster presentation at AcademyHealth, Boston, MA, June 29, 2010 and American Heart Association Quality of Care and Outcomes Research, Scientific Sessions, Washington, DC, May 13, 2011; oral presentation at the Society of General Internal Medicine, Phoenix, AZ, May 5, 2011.

The views expressed in this article are those of the authors and do not necessarily represent the views of the National Institutes of Health, Boston University or Department of Veterans Affairs.

Reprints: Amresh D. Hanchate, PhD, Health/Care Disparities Research Program, Section of General Internal Medicine, Boston University School of Medicine, 801 Massachusetts Ave, #2092, Boston, MA 02118. E-mail: hanchate@bu.edu.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website (www.lww-medicalcare.com).

ISSN: 0025-7079/12/5007-0569

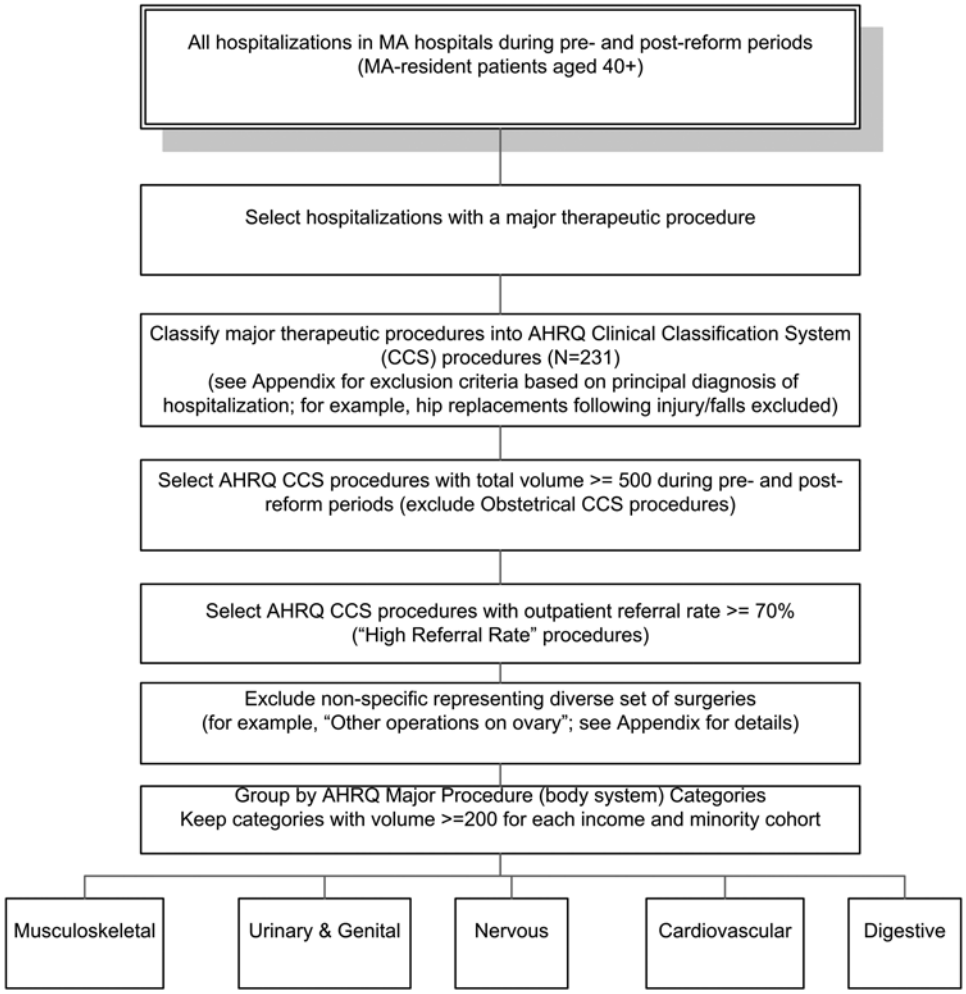


FIGURE 1. Identification and grouping of high referral rate procedures. CCS indicates Clinical Classifications Software; MA, Massachusetts.

improved postreform; however, lower income respondents and Hispanics with limited English proficiency reported higher rates of unmet need due to difficulty in finding a health care provider or due to unaffordable cost.[12–14] Although the evidence of the impact of health reform on emergency department use is mixed, hospitalizations for conditions preventable by appropriate outpatient care decreased.[15–18]

There have not yet been reports on the use of inpatient surgical procedures whose receipt is sensitive to outpatient physician referral. Changes in receipt of such surgical procedures after MA health reform could be a measure of access to outpatient care that may improve with expanded insurance coverage.[19] Thus, we focused on the use of such procedures among vulnerable subpopulations—those living in low-income areas and racial/ethnic minorities. These groups are known to underutilize elective inpatient care,[20–22] and were specifically targeted for larger gains in coverage expansion from the MA health reform.[23] We hypothesized that the entire reform package, including increased population rates of insurance coverage, would be associated with increased rates of receipt of surgical procedures most commonly initiated through outpatient physician referral, and that such gains would be greater among vulnerable populations.

METHODS

Overview

We estimated longitudinal population rates of receipt of referral-dependent procedures by combining comprehensive state-level inpatient administrative data with census population data. We estimated prereform and postreform rates of procedure use among nonelderly subpopulations stratified by cohorts defined by race/ethnicity and income of the area of patient residence. To isolate the impact of health reform from secular trends, we contrasted postreform change among the nonelderly with changes among the elderly. We chose this control group because most elderly residents were covered by Medicare both prereform and postreform, and therefore the change in their procedure use reflects secular changes unrelated to health reform.

Prereform and Postreform Periods

The MA health reform was multifaceted and included measures to expand insurance coverage, such as individual and employer mandates, establishment of an insurance exchange, income-related premium subsidies for newly created private insurance, and loosened eligibility criteria for Medicaid coverage.[2] Implementation of MA health reform began on 7/1/2006 with expansion of Medicaid to cover previously "enrollment capped" low-income populations, culminating in a penalty-enforced mandate of individual insurance coverage effective 1/1/2008.[24] We examined inpatient procedure use for 21 months (1/1/2008–9/30/2009) following this mandate ("postreform" period) and contrasted it with data for 21 months (10/1/2004–6/30/2006) preceding reform ("prereform" period), excluding the middle, transition period.

Data Sources and High Referral Rate (HRR) Procedures

We focused on hospitalizations for surgical procedures that are predominantly scheduled by outpatient referral, that we term "HRR procedures." While similar to the previously defined concept of "referral-sensitive procedures,"[19] we found that for some referral-sensitive procedures (eg, coronary artery bypass graft) the proportion arising from outpatient referral was no higher than 50% (See Appendix, Supplemental Digital Content, http://links.lww.com/MLR/A311). Using all hospitalizations with discharges during the prereform and postreform periods as raw data (MA Inpatient Discharge Data for 2004–2009),[25] we included MA-residing patients aged 40 or older (those with significant risk for the procedures examined) undergoing a major therapeutic surgical procedure [using Agency for Healthcare Research & Quality Procedure Classes system; Clinical Classifications Software (CCS)], as illustrated by Figure 1.[26] These procedures' International Classification of Diseases-Ninth Edition-Clinical Modification diagnosis codes were classified into the 231-category Agency for Healthcare Research & Quality CCS.[27] To minimize chance misclassification we only included procedures with aggregate volume of ≥ 500. We excluded obstetrical procedures as their usage has been universally covered in MA.

On the basis of the "source of admission" field, we defined HRR procedures as those with an outpatient physician referral rate $\geq 70\%$, reasoning that this threshold would represent a large majority of procedures. We excluded some nonspecific HRR procedures—for instance, "Other operations of the ovary"—that captured a broad range of procedures. To minimize chance fluctuations in procedure use, we grouped the HRR CCS procedures into International Classification of Diseases-Ninth Edition-Clinical Modification procedure categories[27] and excluded those with ≤ 200 surgeries for each area income or race/ethnicity cohort.

Information on patient race/ethnicity was part of the discharge data submitted by each hospital; as such identification is likely based on multiple sources including, patient self-report and administrative records. We found longitudinal consistency in the reporting patterns over years, not only for the main racial/ethnic groups (whites, blacks, and Hispanics), but also for proportion with missing race/ethnicity; the proportion of all discharges in a year with race/ethnicity missing or "other" (ie, not white, black, or Hispanic) ranged between a low of 5.32% (2004) to 6.06% (2006) of the study period years (2004–2006, 2008–2009).

Analytic Data Structure

To estimate prereform and postreform procedure use we produced 2 analytic datasets, 1 for performing comparisons by race and ethnicity and another for comparisons by area income. The first dataset was obtained by stratifying the state population into cohorts stratified by race/ethnicity, age, sex, county, and time period (ie, prereform/postreform). We stratified patients by the 3 largest race/ethnic cohorts: Hispanics, non-Hispanic whites, and non-Hispanic blacks (See Appendix, Supplemental Digital Content, http://links.lww.com/MLR/A311). Categorizing patient age (in years) into 10 five-year age groups aged 40–84 years (eg, 40–44 y), we excluded the 65–69-year age group; as the postreform study period lasts 21 months, inclusion of both 60–64- and 65–69-year age cohorts may overestimate reform effect on procedure use if those initially aged 63 or 64 then age into the 65–69-year age group and became eligible for Medicare before the end of the 21st month. Excluding this age group also eliminates the sharp increases in procedure use previously noted for new Medicare age-65 enrollees.[28,29] To adjust for geographic heterogeneity across MA in factors determining procedure use, we stratified the state into 11 county-based areas, as this is the finest substate level for which annual census population counts are available.[30] This allowed us to perform a finer grained analysis than that at the larger state level. With each county stratified into 54 cohorts (based on sex, 9 age groups and 3 racial/ethnic groups), there was a total of 594 observations each for the prereform and postreform periods ($N = 1188$).

For the second analytic dataset, we followed a similar process but replaced race/ethnicity strata by area income strata. In the absence of individual income, we followed previous work and used the median income (2000 census) for each patient's residence zip code to stratify all patients into 3 area income groups: lowest quartile (ie, all residents of zip codes in the poorest quartile, henceforth referred to as "low area income" population), second lowest quartile (medium area income), and top 2 quartiles combined (high area income).[22,31–33] As the number of area income cohorts ($N = 3$) is the same as that number of race/ethnicity cohorts, the second analytic dataset has the same number of observational units ($N = 1188$).

Procedure Rates

Our primary outcome measure was a procedure rate for each cohort of interest, derived from the ratio of (1) the number of HRR procedures for each cohort in the inpatient discharge data; and (2) the census population of this cohort, and then multiplying this ratio by 10,000 so as to obtain the procedure rate per 10,000 census population.

Analysis

We estimated prereform and postreform procedure rates for all HRR procedures combined and for each individual procedure category, for subpopulations by area income or race/ethnicity; we adjusted for compositional

TABLE 1. High Referral Rate Procedures: Volumes and Referral Rate

Procedure Category	# Procedures in Study Period	Average Referral Rate (%)	# Individual Procedures in Category	Individual Procedures (% Share of Category Volume)
Musculoskeletal	80,688	95	4	Knee arthroplasty (40%), spinal fusion (31%), hip replacement (22%), and partial excision bone (7%)
Urinary/genital	51,088	95	6	Hysterectomy (32%), oophorectomy (32%), repair of cystocele and rectocele (12%), transurethral resection of prostate (12%), genitourinary incontinence procedures (8%), and nephrectomy (2%)
Nervous	29,372	88	1	Laminectomy (100%)
Cardiovascular	20,914	80	4	Heart valve procedures (34%), endarterectomy (32%), peripheral vascular bypass (19%), and aortic resection (15%)
Digestive	19,845	73	2	Colorectal resection (98%) and gastrectomy (2%)
All procedures	201,907	90	17	

Procedures are categorized by ICD-9-CM procedure chapters. No eye, ear, and nose/throat procedures met the high referral rate criteria. Urinary, male genital, and female genital system procedures have been combined into 1 category.

Study period comprises prereform and postreform periods (10/1/2004–6/30/2006 and 1/1/2008–9/30/2009, respectively).

procedures indicates the state-wide volume performed during hospitalization stays in the prereform and postreform periods.

Average referral rate indicates the % of hospitalizations (during which procedure was performed) based on outpatient referral.

differences in sex and age by direct standardization.[34] These adjusted rates were estimated separately for nonelderly and elderly cohorts. We first measured overall change (%) in procedure use—the percentage change between prereform and postreform procedure rates. To estimate net change (%) associated with health reform, we adjusted for secular changes using the elderly as the comparison cohort ("difference-in-difference" estimation).[35,36] We used a count regression (Poisson) model with procedure count as the outcome measure and census population count as the population at risk. We specified a county-level fixed effects regression structure (with clustering-adjusted SEs) to capture nesting of cohorts within county.[35,37] Regression covariates included indicators of age, sex, race/ethnicity or area income, time period (prereform/postreform), and interaction between the elderly/nonelderly indicator and time period (to estimate the net change). Statistical significance was assessed at the level of $P < 0.05$. All estimation was performed using Stata Version 11.1.[37] We performed several sensitivity analyses to assess the robustness of findings to (a) inclusion of 65–69 age group; (b) alternative count regression (ie, negative binomial) specification; and (c) state-level aggregated unit of analysis (ie, without county-level stratification). To address potential bias from regression to the mean or differential changes in the characteristics of the study over time, we also estimated an alternative model on the basis of segmented time-series specification of postreform effects that allowed for level and trend effects. This study was approved by the Boston University Medical Campus Institutional Review Board.

Note that the comparison groups were based on age, not whether the patient held Medicare coverage, so dual eligibles were included with their respective age groups. Our estimates are based on change in procedure rate (say, among nonelderly or elderly patients) between the prereform and postreform period. It may well be that the subgroup of, say, elderly with dual eligibility may have different procedure rates than their counterparts without dual eligibility; however, to the extent that prevalence of dual eligibility remained similar in the prereform and postreform periods, it does not impact the net estimates that we have estimated.

RESULTS

We identified 17 HRR procedures, in 5 clinical categories, with an aggregate volume of 201,907 surgeries during the prereform and postreform periods (Table 1). The overall referral rate for all procedures was 90%.

Table 2 summarizes the number of people undergoing HRR procedures, the number of people in the population at risk, and each group's sociodemographic composition prereform and postreform. Whereas the nonelderly accounted for 60% of prereform surgeries, their share increased to 64% in the postreform period; however, the nonelderly share of the population at risk remained at 78%. The share of blacks and Hispanics increased both among procedure recipients (6.4%–8.2%) and the population at risk (9.0%–9.8%); share by area income cohorts did not change.

Procedure Rates and Overall Postreform Change by Area Income

Prereform use of HRR procedures was similar among nonelderly area income subgroups (Table 3). After reform, overall increases in procedure rates were higher for low area income (8%, $P = 0.04$) and medium area income (10%, $P < 0.001$) cohorts compared with that for their high area income counterparts (4%). Adjusting for secular changes, the impact of health reform for the nonelderly income cohorts (or the net change in procedure rate) was: 13% (low area income, 95% confidence intervals (CI) = [9%, 17%]), 15% (medium area income, 95% CI = [6%, 24%]), and 2% (high area income, 95% CI = [−3%, 8%]).

Procedure Rates and Overall Postreform Change by Race/Ethnicity

Prereform use of all HRR procedures was significantly lower among nonelderly Hispanics (118 procedures per 10,000 population; $P < 0.001$) and blacks (149; $P = 0.05$) compared with whites (157). After reform, overall change in procedure rates among the nonelderly was greater among Hispanics (23%, $P < 0.001$) and blacks (21%, $P < 0.001$) compared with that among whites (7%). Adjusting for secular trends, the net change in procedure rate was 22% (95%

TABLE 2. Counts of High Referral Rate Procedures and Population at Risk, by Sociodemographics

Characteristics	Total # Procedures (Prereform and Postreform Periods)	Distribution of Procedures		# Population at Risk, Person-Years (Prereform and Postreform Periods)	Distribution of Population at Risk	
		Prereform (%)	Postreform (%)		Prereform (%)	Postreform (%)
All (age 40+ y)	201,907	100	100	10,072,992	100	100
Women	123,560	60	62	5,387,207	54	53
Men	78,347	40	38	4,685,785	46	47
Age (y)						
40–64	124,966	60	64	7,845,648	78	78
70 or older	76,941	40	36	2,227,344	22	22
Area income, zip code median						
Low	43,880	22	22	2,282,028	23	23
Medium	50,454	25	25	2,494,057	25	25
High	107,229	53	53	5,296,841	53	53
Race/ethnicity						
Hispanics	6688	2.8	3.8	484,209	4.5	5.1
Blacks	8115	3.6	4.4	460,728	4.5	4.7
Others	6324	3.4	2.8	419,872	3.9	4.4
Whites	180,780	90.1	89.0	8,708,183	87.1	85.8

Study period comprises prereform and postreform periods (10/1/2004–6/30/2006 and 1/1/2008–9/30/2009, respectively).
Total # procedures indicates the number of high referral rate procedures performed during hospitalizations in the prereform and postreform periods.
Population at risk indicates the aggregate census population (person-years) during the prereform and postreform periods.
Area income cohorts are defined as: low = lowest quartile zip codes, medium = second quartile, and high = top 2 quartile.

CI = [5%, 38%]) for Hispanics, 5% for blacks (95% CI = [−20%, 30%]), and 7% for whites (95% CI = [5%, 10%]).

Postreform Change by Procedure Categories

Table 4 presents analogous findings for each procedure category by area income and race/ethnicity. There is considerable variation in overall and net changes across categories, with some indicating secular decrease in procedure rates, but statistical precision of estimates is also reduced because of relatively smaller volumes within individual procedure categories. For musculoskeletal and urinary/genital procedures, both low and medium area income cohorts experienced significant increase in overall postreform procedure rates. Although not statistically significant, we found that compared with the high area income group, the estimated net increase (%) was larger or net decrease smaller for all 5 procedure categories among the lowest area income cohort and for 4 procedure categories among the medium area income cohort. Comparisons by race/ethnicity indicate that compared with whites, the estimated net change (%) was greater for 3 of the procedure types (musculoskeletal, urinary/genital, and nervous) among Hispanics, but only for digestive procedures among blacks—and statistically significant only for nervous system procedures among Hispanics.

Sensitivity analyses indicate that all main findings reported are robust to (a) inclusion of those aged 65–69; (b) alternative count regression specification to permit overdispersion; and (c) aggregation of procedure counts to state instead of county level (See Appendix, Supplemental Digital Content, http://links.lww.com/MLR/A311). Segmented time-series Poisson model indicated similar patterns in postreform change, with no significant transition period effects for any of the cohorts (See Appendix, Supplemental Digital Content, http://links.lww.com/MLR/A311).

DISCUSSION

We compared prereform and postreform utilization of major therapeutic inpatient surgical procedures predominantly scheduled by outpatient referrals among nonelderly MA adults, and found greater overall increases for lower area income cohorts compared with the highest area income cohort, and for Hispanics compared with whites. Before reform, both blacks and Hispanics had lower rates of these procedures compared with whites. We estimated the net change in procedure use associated with health reform among the nonelderly accounting for secular trends, finding significant increases for lower area income groups and Hispanics and whites but not among blacks or the highest area income group. As 90% of all surgeries came from outpatient physician referral, these findings suggest a meaningful improvement in access to outpatient care for the surgeries studied, especially those living in lower income areas, Hispanics, and whites.

Our findings of greater net increases in procedure use among lower area income groups and Hispanics are consistent with previous randomized[38,39] and natural experiments of expanded public insurance programs or similar policy changes; however, few prior studies have explicitly examined whether increased insurance coverage reduces income or racial/ethnic disparities in access to or use of care.[36,40] A recent study of Oregon's lottery-selected expansion of Medicaid to uninsured low-income nonelderly

TABLE 3. Prereform and Postreform Use of High Referral Rate Procedure by Area Income and Race/Ethnicity Cohorts

	Prereform Procedure Rate		Postreform Procedure Rate			Excess Overall Change among Minorities/Lower Area Income Cohorts		Net Change (%): Postreform Change (%) in Procedure Rate Among Nonelderly Because of Health Reform	
Cohorts	# Procedures/ 10,000	95% Confidence Interval	# Procedures/ 10,000	95% Confidence Interval	Overall Postreform Change in Procedure Rate (%)	Difference in Change Between Nonwhite/White or Lower Area Income/Highest Income (%)	*P* (Difference = 0)	Net Change (%)	95% Confidence Interval
By area income cohorts									
Nonelderly (age 40–64 y)									
Low	156	[154%, 159%]	169	[167%, 172%]	8	4	0.04	13	[9%, 17%]
Medium	151	[149%, 154%]	166	[164%, 169%]	10	6	<0.001	15	[6%, 24%]
High	154	[153%, 156%]	161	[159%, 162%]	4	Reference		2	[−3%, 8%]
Elderly (age 70 or older)									
Low	302	[295%, 308%]	284	[278%, 291%]	−6	−7	<0.001		
Medium	354	[347%, 361%]	340	[333%, 346%]	−4	−5	0.002		
High	365	[360%, 370%]	369	[365%, 374%]	1	Reference			
By race/ethnicity cohorts									
Nonelderly (age 40–64 y)									
Hispanics	118	[113%, 123%]	146	[141%, 151%]	23	16	<0.001	22	[5%, 38%]
Blacks	149	[143%, 155%]	180	[174%, 186%]	21	14	<0.001	5	[−20%, 30%]
Whites	157	[155%, 158%]	167	[166%, 168%]	7	Reference		7	[5%, 10%]
Elderly (age 70 or older)									
Hispanics	199	[181%, 216%]	202	[186%, 218%]	2	3	0.926		
Blacks	234	[218%, 251%]	268	[251%, 285%]	14	15	0.003		
Whites	355	[352%, 359%]	353	[349%, 356%]	−1	Reference			

Prereform = [October 2004, June 2006] and postreform = [January 2008, September 2009].

Procedure rates, adjusted for sex and age differences, are estimated by direct standardization method.

Overall change (%) in procedure rate is the simple % change in the prereform and postreform procedure rates. Difference in overall change by race/ethnicity and area income are obtained by comparing overall change.

Net change (%) in procedure rate among nonelderly due to health reform are calculated from a separate (county fixed effects) Poisson regression with a difference-in-difference specification using combined nonelderly and elderly cohort data. Confidence intervals are based on bootstrapped SEs (N = 1000 iterations) to adjust for clustering of observations within county.

Area income cohorts are defined as: low = lowest quartile zip codes, medium = second quartile, and high = top 2 quartiles.

adults in 2008 found that hospital admissions increased by 30% in 1 year; this effect is nearly identical to that found in the RAND randomized study in the 1970s.[38,39] More relevant to our study is the finding from Oregon that the increase in inpatient admissions was "disproportionately concentrated" among admissions "that do not originate in the emergency room"; we note that these primarily include admissions based on outpatient physician referral, including those for HRR procedures examined here.[39]

More appropriate for comparison with our study are findings of quasi-experimental expansions of public health insurance.[28,40–42] Studies examining the impact of Medicare enrollment at age 65 have noted increased use of inpatient and outpatient care among the previously uninsured[29] and also the previously insured (because of the relative "generosity" of Medicare).[28,40] One study documented a 10% increase in hospitalizations in the year after Medicare enrollment, with larger increases in use of "elective" procedures such as bypass surgery and joint replacement.[28] This suggests that our finding of increased procedure use may reflect a combination of pent-up unmet need and need arising from new diagnoses after increased access to outpatient care.

Although the 17 surgical procedures examined represent a broad spectrum of inpatient procedures, our main focus here was on their role as markers of access to care. In combining these procedures for evaluating the differential impact of health reform in access to care across subpopulations, we recognize heterogeneity in the procedures in other respects, including acuity of conditions targeted, impact on quality of life, and value in terms of clinical benefit per dollar. Reflecting this heterogeneity, we found considerable differences in postreform changes in rates, with several categories of procedures experiencing decrease in utilization while some others had sharp increases ($\geq$ 25%). As estimates of net increases by individual procedure categories had wide CIs due largely to small numbers, we cannot rule out potentially large differences among subpopulations. Nevertheless, statistically significant net increases associated with health reform were found for musculoskeletal and urinary/genital procedures among lower area income cohorts and whites, and for urinary/genital procedures among Hispanics.

TABLE 4. Impact of Health Reform on Use of High Referral Rate Procedures by *Area* Income, Race/Ethnicity & Procedure Category

		HRR Procedure Rates Nonelderly (40–64 y)		HRR Procedure Rates Elderly (70 or Older)		Net change (%): Postreform Change in Procedure Rate Among Nonelderly due to Health Reform			HRR Procedure Rates Nonelderly (40–64 y)		HRR Procedure Rates Elderly (70 y or Older)		Net Change (%): Postreform Change in Procedure Rate Among Nonelderly due to Health Reform	
Procedure Category	Cohort (Area Income)	Overall Postreform Change in Procedure Rate (%)	*P* (Difference in Lower Vs. Highest Area Income Cohorts % Change = 0)	Overall Postreform Change in Procedure Rate (%)	*P* (Difference in Lower Vs. Highest Area Income Cohorts % Change = 0)	Net Change (%)	95% Confidence Interval	Cohort (Race/ Ethnicity)	Overall Postreform Change in Procedure Rate (%)	*P* (Difference in Nonwhite Vs. White % Change = 0)	Overall Postreform Change in Procedure Rate (%)	*P* (Difference in Nonwhite Vs. White % Change = 0)	Net Change (%)	95% Confidence Interval
Musculoskeletal														
	Low	20	<0.001	5	<0.001	12	[3%, 22%]	Hispanics	31	0.017	19	0.354	8	[−4%, 31%]
	Medium	18	0.001	5	0.001	12	[5%, 19%]	Blacks	25	0.028	20	0.124	3	[−12%, 19%]
	High	9	Reference	9	Reference	1	[−3%, 6%]	Whites	15	Reference	8	Reference	6	[2%, 11%]
Urinary/genital														
	Low	12	0.595	−11	0.397	21	[10%, 31%]	Hispanics	30	<0.001	−1	0.846	34	[1%, 67%]
	Medium	16	0.031	−15	0.024	32	[21%, 42%]	Blacks	40	<0.001	42	<0.001	1	[−22%, 24%]
	High	9	Reference	−7	Reference	15	[2%, 28%]	Whites	10	Reference	−10	Reference	19	[11%, 26%]
Nervous														
	Low	−6	0.921	−2	0.034	−4	[−16%, 9%]	Hispanics	8	0.151	−47	<0.001	71	[2%, 140%]
	Medium	−3	0.32	3	0.208	−6	[−20%, 8%]	Blacks	−16	0.122	43	0.03	−51	[−127%, 25%]
	High	−7	Reference	10	Reference	−15	[−28%, −3%]	Whites	−3	Reference	7	Reference	−10	[−19%, −2%]
Cardiovascular														
	Low	−19	0.312	−10	0.157	−12	[−19%, −4%]	Hispanics	−2	0.265	15	0.216	−14	[−52%, 24%]
	Medium	−15	0.849	−13	0.021	5	[−6%, 16%]	Blacks	2	0.106	4	0.5	0	[−65%, 66%]
	High	−15	Reference	−4	Reference	−16	[−21%, 11%]	Whites	−17	Reference	−7	Reference	−11	[−15%, −7%]
Digestive														
	Low	−5	0.989	−24	0.005	23	[9%, 37%]	Hispanics	8	0.245	4	0.508	10	[−27%, 48%]
	Medium	−4	0.865	−9	0.505	0	[−10%, 11%]	Blacks	5	0.375	−30	0.153	37	[5%, 69%]
	High	−5	Reference	−12	Reference	10	[2%, 19%]	Whites	−4	Reference	−13	Reference	10	[5%, 15%]

Overall change (%) in procedure rate is the simple % change in the prereform and postreform procedure rates. Difference in overall change by race/ethnicity and area income are obtained by comparing overall change. The figures for the all cohort are identical to those in Table 3.

Net change (%) in procedure rate among nonelderly due to health reform are calculated from a separate (county fixed effects) Poisson regression with a difference-in-difference specification using combined nonelderly and elderly cohort data. Confidence intervals are based on bootstrapped SEs (N = 1000 iterations) to adjust for clustering of observations within county.

Area income cohorts are defined as: low = lowest quartile zip codes, medium = second quartile, and high = top 2 quartiles.

HRR indicates high referral rate.

For Hispanics, the overall postreform increase in procedure use among the nonelderly was considerably higher than that for their elderly counterparts, particularly for musculoskeletal, urinary/genital, and nervous system procedures. For blacks, whereas the changes for both groups were similar for musculoskeletal and urinary/genital procedures, the magnitude of the change is large and comparable with that for the nonelderly Hispanics. Therefore, it is the similar increase in the use of these procedures among the elderly blacks that leads to the results of no significant net change (for nonelderly) attributable to the reform. Reasons for the similar increase among all blacks (elderly and nonelderly) are unclear and merit further examination.

There is considerable debate on whether more medical care leads to better health.[43] However, most studies of natural experimental policy changes have found that expansions of health insurance result in health improvements for individual health measures or subpopulations.[36] Given the natural experimental setting of MA reform, we instead examined disparities in health care utilization and focused on vulnerable subpopulations and selected inpatient procedure categories for which underutilization of care is known to be associated with uninsurance or underinsurance. Research has documented higher rates of clinically unmet need among minorities and lower income patients for many inpatient procedures, including those for cardiac,[44] cancer,[45,46] and musculoskeletal[22] care. Our findings are among the first to show that expanded insurance coverage on a population level is associated with increase in use by such vulnerable populations.

Our study has several important limitations. First, we cannot differentiate overuse of procedures from clinically appropriate use. We suspect that our findings of increased procedure use among minorities do not reflect overuse, as Dartmouth Atlas comparisons of regional differences for Medicare beneficiaries for 12 common inpatient surgeries found MA procedure rates were below average for 6, near average for 5, and above average for only 1 procedure.[47] Second, as our data is observational, the possibility of potential confounding from unobserved factors remains. However, as we adjust for changes among the elderly, our estimates are robust to unobserved factors (including practice pattern changes) that affect all age groups. Also, comparison of nonelderly and elderly rates of use may not be clinically meaningful for some procedures. However, our findings do include same-age group comparisons by race/ethnicity and area income cohorts. Further, we did not include individual-level data on insurance status, because of the inability to infer population rates of insurance status by the subgroups of interest from our data on health care users only. Identification of patient race/ethnicity is not necessarily based on patient self-report and may vary across hospitals; however, as this is likely to affect both nonelderly and elderly patients in each hospital, our methodology of contrasting changes among nonelderly patients with those for elderly patients provides robustness of findings to the potential heterogeneity in race/ethnicity identification. Also, in the absence of data on individual income, we have used zip code–level income as the measure of socioeconomic status; however, this approach has been used in numerous previous studies.[31–33] Finally, our focus on the use of inpatient procedures may underestimate use of procedures performed in outpatient settings.

Nonetheless our findings have implications for national health reform (Affordable Care Act, 2010) which shares many key elements with MA health reform.[1] Notably, before health reform, MA had lower uninsurance and better safety-net funding compared with other states.[48,49] Depending on the extent to which similar subpopulations gain from insurance expansion from the national reform, the potential for improved access is considerably larger or smaller, as is the potential for higher costs. Our study examined utilization only in the first 2 years after the reform, and therefore may include sharp increases in utilization from nonelderly patients with prior unmet need. Whether these increases will taper-off in the longer run is unknown. Actual changes also depend on other factors, including provider supply and practice patterns, which also vary considerably across states.

In conclusion, our findings of significant postreform expansion in procedure use for Hispanics and lower area income patients are consistent with the relatively larger gains in insurance coverage among these subpopulations. These findings suggest potentially improved access to outpatient care and may reflect demand built up before reform when individuals were uninsured. Whether such improved access—a crucially important first step to improving equity in access and outcomes—translates into improved clinical outcomes at a reasonable cost merits further study.

ACKNOWLEDGMENT

The authors are indebted to Rachel Werner, MD, PhD, for her thoughtful comments on an earlier version of this manuscript.

REFERENCES

1. Davis K. *A New Era in American Health Care: Realizing the Potential of Reform*. New York, NY: The Commonwealth Fund; 2010.
2. Holahan J, Blumberg L. Massachusetts health care reform: a look at the issues. *Health Aff (Millwood)*. 2006;25:w432–w443.
3. Kingsdale J. Implementing health care reform in Massachusetts: strategic lessons learned. *Health Aff (Millwood)*. 2009;28:w588–w594.
4. Turnbull NC. The Massachusetts model: an artful balance. *Health Aff (Millwood)*. 2006;25:w453–w456.
5. Weissman JS, Bigby J. Massachusetts health care reform—near-universal coverage at what cost? *N Engl J Med*. 2009;361:2012–2015.
6. Wilson JF. Massachusetts health care reform is a pioneer effort, but complications remain. *Ann Intern Med*. 2008;148:489–492.
7. Gruber J. Massachusetts points the way to successful health care reform. *J Policy Anal Manag*. 2011;30:184–192.
8. MA Division of Health Care Finance and Policy. *Health Care in Massachusetts: Key Indicators*. Boston, MA: Division of Health Care Finance and Policy; 2011.
9. Long SK, Phadera L. *Estimates of Health Insurance Coverage in Massachusetts From the 2009 Massachusetts Health Insurance Survey*. Washington DC: Urban Institute; 2009.
10. McDonough JE, Miller M, Barber CA. Progress report on state health access reform. *Health Aff*. 2008;27:w105–w115.
11. Massachusetts Department of Public Health. *A Profile of Health Among Massachusetts Adults, 2006 & 2009. Results from the Behavioral Risk Factor Surveillance System*. Boston: Commonwealth of Massachusetts; 2010.

12. Long SK. On the road to universal coverage: impacts of reform in Massachusetts at one year. *Health Aff*. 2008;27:w270–w284.
13. Zhu J, Brawarsky P, Lipsitz S, et al. Massachusetts health reform and disparities in coverage, access and health status. *J Gen Intern Med*. 2010;25:1356–1362.
14. Maxwell J, Cortés DE, Schneider KL, et al. Massachusetts' health care reform increased access to care for Hispanics, but disparities remain. *Health Aff*. 2011;30:1451–1460.
15. Kolstad JT, Kowalski AE. The Impact of an Individual Health Insurance Mandate On Hospital and Preventive Care: Evidence from Massachusetts. National Bureau of Economic Research Working Paper Series 2010;No. 16012.
16. MA Division of Health Care Finance and Policy. *Preventable/Avoidable Emergency Department Use in Massachusetts Fiscal Years 2004 to 2008*. Boston: Division of Health Care Finance and Policy, Massachusetts Executive Office of Health and Human Services; 2010.
17. MA Division of Health Care Finance and Policy. *Potentially Preventable Hospitalizations In Massachusetts Fiscal Years 2004–2008*. Boston: Division of Health Care Finance and Policy, Massachusetts Executive Office of Health and Human Services; 2010.
18. Miller S. *The Effect of Insurance on Outpatient Emergency Room Visits: An Analysis of the 2006 Massachusetts Health Reform. Urbana-Champaign*. IL: University of Illinois; 2011.
19. Billings J. Using administrative data to monitor access, identify disparities, and assess performance of the safety net. In: Weinick RM, Billings J, eds. *Tools for Monitoring the Health Care Safety Net*. Rockville, MD: Agency for Healthcare Research and Quality; 2003.
20. Kressin NR, Petersen LA. Racial differences in the use of invasive cardiovascular procedures: review of the literature and prescription for future research. *Ann Intern Med*. 2001;135:352–366.
21. Ward E, Jemal A, Cokkinides V, et al. Cancer disparities by race/ethnicity and socioeconomic status. *CA Cancer J Clin*. 2004;54:78–93.
22. Skinner J, Zhou W, Weinstein J. The influence of income and race on total knee arthroplasty in the United States. *J Bone Joint Surg Am*. 2006;88:2159–2166.
23. McDonough JE, Rosman B, Phelps F, et al. The third wave of Massachusetts health care access reform. *Health Aff (Millwood)*. 2006;25:w420–w431.
24. McDonough JE, Rosman B, Butt M, et al. Massachusetts health reform implementation: major progress and future challenges. *Health Aff (Millwood)*. 2008;27:w285–w297.
25. MA Division of Health Care Finance and Policy. *FY2007 Inpatient Hospital Discharge Database Documentation Manual*. Boston: Division of Health Care Finance and Policy, Massachusetts Executive Office of Health and Human Services; 2009.
26. Agency for Healthcare Research and Quality. Procedure Classes 2011. Rockville, MD: Department of Health and Human Services, Agency for Healthcare Research and Quality; 2010.
27. Agency for Healthcare Research and Quality. Clinical Classifications Software (CCS) 2010. Rockville, MD: Department of Health and Human Services, Agency for Healthcare Research and Quality; 2010.
28. Card D, Dobkin C, Maestas N. The impact of nearly universal insurance coverage on health care utilization and health: Evidence from Medicare. NBER Working paper No. 10365. Cambridge, MA; 2004.
29. McWilliams JM, Meara E, Zaslavsky AM, et al. Use of health services by previously uninsured Medicare beneficiaries. *N Engl J Med*. 2007; 357:143–153.
30. Census Bureau. Methodology for the State and County Resident Population Estimates by Age, Sex, Race and Hispanic Origin (Vintage 2008): April 1, 2000 to July 1, 2008. 2010.
31. Jha AK, Fisher ES, Li Z, et al. Racial trends in the use of major procedures among the elderly. *N Engl J Med*. 2005;353:683–691.
32. Chernew M, Gibson TB, Yu-Isenberg K, et al. Effects of Increased Patient Cost Sharing on Socioeconomic Disparities in Health Care. *J Gen Intern Med*. 2008;23:1131–1136.
33. Fiscella K, Franks P. Impact of patient socioeconomic status on physician profiles: a comparison of census-derived and individual measures. *Med Care*. 2001;39:8–14.
34. Woodward M. *Epidemiology: Study Design and Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC; 2005.
35. Cameron AC, Trivedi PK. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press; 2005.
36. Levy H, Meltzer D. The impact of health insurance on health. *Annu Rev Public Health*. 2008;29:399–409.
37. StataCorp. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP; 2009.
38. Manning WG, Newhouse JP, Duan N, et al. Health insurance and the demand for medical care: evidence from a randomized experiment. *Am Econ Rev*. 1987;77:251–277.
39. Finkelstein A, Taubman S, Wright B, et al. The Oregon Health Insurance Experiment: Evidence from the First Year. National Bureau of Economic Research Working Paper Series 2011;No. 17190.
40. Finkelstein A. The aggregate effects of health insurance: evidence from the introduction of Medicare. *Q J Econ*. 2007;122:1–37.
41. Buchmueller TC, Grumbach K, Kronick R, et al. The effect of health insurance on medical care utilization and implications for insurance expansion: a review of the literature. *Med Care Res Rev*. 2005;62:3–30.
42. Finkelstein A, McKnight R. What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending. *J Public Econ*. 2007;92:1644–1669.
43. Fisher E. Medical care—is more always better? *N Engl J Med*. 2003; 349:1665–1667.
44. Rao SV, Schulman KA, Curtis LH, et al. Socioeconomic status and outcome following acute myocardial infarction in elderly patients. *Arch Intern Med*. 2004;164:1128–1133.
45. Ayanian JZ, Kohler BA, Abe T, et al. The relation between health insurance coverage and clinical outcomes among women with breast cancer. *N Engl J Med*. 1993;329:326–331.
46. Roetzheim RG, Pal N, Tennant C, et al. Effects of health insurance and race on early detection of cancer. *J Natl Cancer Inst*. 1999;91:1409–1415.
47. Dartmouth Medical School. *The Dartmouth Atlas of Health Care in the United States*. Hanover, NH: Center for the Evaluative Clinical Sciences, Dartmouth Medical School; 1999.
48. Kaiser Family Foundation. State Health Facts. http://www.statehealthfacts.org; 2010. Accessed July 15, 2011.
49. Gruber J. The Impacts of the Affordable Care Act: How Reasonable Are the Projections? National Bureau of Economic Research Working Paper Series 2011; No. 17168.

Comparison of Ischemic Stroke Outcomes and Patient and Hospital Characteristics by Race/Ethnicity and Socioeconomic Status

Amresh D. Hanchate, PhD; Lee H. Schwamm, MD, FAHA; Wei Huang, MPH, MA; Elaine M. Hylek, MD, MPH

Background and Purpose—Current literature provides mixed evidence on disparities by race/ethnicity and socioeconomic status in discharge outcomes after hospitalization for acute ischemic stroke. Using comprehensive data from 8 states, we sought to compare inpatient mortality and length of stay by race/ethnicity and socioeconomic status.

Methods—We examined all 2007 hospitalizations for acute ischemic stroke in all nonfederal acute care hospitals in Arizona, California, Florida, Maine, New Jersey, New York, Pennsylvania, and Texas. Population was stratified by race/ethnicity (non-Hispanic whites, non-Hispanic blacks, and Hispanics) and socioeconomic status, measured by median income of patient zip code. For each stratum, we estimated risk-adjusted rates of inpatient mortality and longer length of stay (greater than median length of stay). We also compared the hospitals where these subpopulations received care.

Results—Hispanic and black patients accounted for 14% and 12% of all ischemic stroke admissions (N=147 780), respectively, and had lower crude inpatient mortality rates (Hispanic=4.5%, blacks=4.4%; all $P<0.001$) compared with white patients (5.8%). Hispanic and black patients were younger and fewer had any form of atrial fibrillation. Adjusted for patient risk, inpatient mortality was similar by race/ethnicity, but was significantly higher for low-income area patients than that for high-income area patients (odds ratio, 1.08; 95% confidence interval, 1.02–1.15). Risk-adjusted rates of longer length of stay were higher among minority and low-income area populations.

Conclusions—Risk-adjusted inpatient mortality was similar among patients by race/ethnicity but higher among patients from lower income areas. However, this pattern was not evident in sensitivity analyses, including the use of mechanical ventilation as a partial surrogate for stroke severity. **(*Stroke*. 2013;44:469-476.)**

Key Words: ethnicity ■ ischemic stroke ■ inpatient mortality ■ length of stay ■ race ■ risk factors ■ socioeconomic status

Although key to understanding the long-term racial/ethnic disparities in stroke mortality rates, evidence on differences in case fatality after stroke has remained mixed.[1,2] Death certificate data have long indicated that the national death rate from stroke is higher (+48% in 2007) among non-Hispanic blacks (henceforth referred to as blacks) and lower (−20% in 2007) among Hispanics than among non-Hispanic whites (henceforth referred to as whites).[3] Higher stroke mortality can arise from higher incidence or higher case fatality or both.[4,5] Findings from several well-known epidemiological studies over the last 2 decades have uniformly indicated that stroke incidence rate is higher among blacks and Hispanics compared with whites[5–11]; however, the magnitude of disparity in stroke incidence is much smaller than that in stroke mortality.[4,5] Although this would suggest higher case fatality rates among blacks compared with whites, previous studies have been largely inconclusive, with studies finding higher, similar, or lower case fatality among blacks compared with whites across different study populations.[4–8,10,12–15] Evidence on differences between Hispanics and whites is also mixed.[12,16] In the absence of a population-representative data source that is both national in scope and has sizable number of minority patients, these differing findings could plausibly be reflective of regional differences, unrepresentative convenience patient cohorts, or inadequate sample size, thereby limiting comparability and generalizability.[4] Of particular importance is the need to accurately capture low-income cohorts in the study population; owing to higher incidence of low incomes among racial and ethnic minorities, they may be underrepresented in survey-based population studies.[4]

With a view to capture a broader national population, including larger numbers of minority populations, we pooled

Received June 27, 2012; final revision received November 11, 2012; accepted November 13, 2012

From the VA Boston Healthcare System, Boston, MA (A.D.H.); Section of General Internal Medicine, Boston University School of Medicine, Boston, MA (A.D.H., E.H.); Department of Neurology, Harvard Medical School and Massachusetts General Hospital, Boston, MA (L.H.S.); Department of Health Policy and Management, Boston University School of Public Health, Boston, MA (W.-J.H.); and Boston Medical Center, Boston, MA (E.H.).

The online-only Data Supplement is available with this article at http://stroke.ahajournals.org/lookup/suppl/doi:10.1161/STROKEAHA.112.669341/-/DC1.

Correspondence to Amresh D. Hanchate, Section of General Internal Medicine, Boston University School of Medicine, 801 Massachusetts Ave, No. 2092, Boston, MA 02118. E-mail hanchate@bu.edu

Stroke **is available at http://stroke.ahajournals.org** **DOI: 10.1161/STROKEAHA.112.669341**

administrative data of all ischemic stroke discharges from 8 states, which together accounted for 73% of national Hispanic and 37% of national black populations. We examined inpatient mortality as the indicator of case fatality; in addition, length of stay (LOS) was also examined as a secondary patient outcome measure.[12,13] We compared outcomes by race/ethnicity and socioeconomic status (SES). We examined for systematic differences in the type of hospitals where minority and low-income area patients received care.

Methods

Data Sources

The primary data sources were the state-level inpatient discharge databases (2007) from Arizona, California, Florida, Maine, New Jersey, New York, Pennsylvania, and Texas. These cover the universe of all admissions to all nonfederal acute care hospitals.[17] These states were selected based on their sizable minority population and completeness of race and ethnicity data; the proportion of acute ischemic stroke admissions with patient race/ethnicity missing (or coded as unknown) was 2.8% and ranged from 0.3% (Texas) to 7.6% (New York). In addition, we obtained zip code-level data on median household income from the 2000 Census.[18] We obtained data on hospital characteristics from the American Hospital Association Annual Survey (2007).[19]

Study Population

To identify acute ischemic stroke admissions for adults ≥18 years of age, we followed other recent studies by including admissions with the principal discharge code (*International Classification of Diseases, Ninth Revision*, Clinically Modified) of 433.x1, 434.x1, and 436.[12,13] To minimize confounding from scheduled admissions for discretionary treatment (for instance, carotid endarterectomy), we only considered inpatient admissions arising either from admission to the emergency department or transfer from another acute care facility. To avoid 2 admissions being associated with the same outcome, we excluded hospitalizations that resulted in transfer to another acute care hospital. For comparability of patient outcomes by subpopulations, especially by SES, we included admissions of patients who were state residents. From the resulting cohort, we excluded 4.9% admissions with missing key measures; compared with the included cohort, patients in this 4.9% excluded cohort were broadly similar (Appendix in the online-only Data Supplement).

Measures

The primary outcome measures were inpatient death and LOS of the index hospitalization for acute ischemic stroke. We examined LOS as a dichotomous measure indicating longer LOS (ie, 1=LOS >median LOS) because it is more robust to outliers and skewed distribution.[12] Patient risk factors were characterized by comorbid conditions captured by the secondary discharge diagnosis codes (*International Classification of Diseases, Ninth Revision*, Clinically Modified).[13,20] We identified patients with the following comorbid conditions: atrial fibrillation, coronary heart disease, congestive heart failure, diabetes mellitus, and hypertension.[21] Indications for other comorbid conditions were collectively grouped using the Charlson comorbidity index.[22] Discharge data from all states included separate indicators of race and ethnicity, which were together used to uniformly define 4 population cohorts: Hispanics, non-Hispanic whites, non-Hispanic blacks, and others; the others included those with missing race/ethnicity.[23] Given the absence of individual measures of SES, we followed previous research and used the census-based median zip code income, matched with patient residence zip code, to stratify patients from each state into 3 groups based on population quartiles[24,25]: lowest quartile zip codes (low-income area), second lowest quartile (medium), and top 2 quartiles (high).[25] On the basis of previous research on the role of hospital characteristics, we examined several structural measures: number of inpatient beds, number of intensive care unit beds, nursing staff (number of patients per nurse), mean daily emergency department admission volume, ownership, teaching status, and safety-net status (defined as >20% of discharges from the hospital covered by Medicaid).[26]

Estimation

The unit of all analyses was the index patient admission for acute ischemic stroke. We performed bivariate comparisons of patient and hospital characteristics by race/ethnicity and SES. We obtained 2 sets of risk-adjusted estimates of differences in the outcomes across population subgroups stratified by race/ethnicity and SES; in the first, we adjusted for patient risk factors, and in the second, we also adjusted for hospital characteristics.[12] Each was estimated by hierarchical logistic regression to adjust for clustering within hospitals. We estimated these models for all patients and for subgroups formed by the interaction of race/ethnicity and SES. Standard errors and confidence intervals (CIs) were estimated to account for heterogeneity across subpopulations. Statistical significance was defined as $P \leq 0.05$.

Across patient risk factors, we found relatively larger differences in age and comorbidity of atrial fibrillation between whites and minorities. To assess their relative impact on outcomes, we estimated racial/ethnic differences in outcome rates adjusted for only these 2 factors.

Sensitivity Analyses

Ideally, we would have liked to adjust for initial stroke severity. In the absence of available data on stroke severity, we performed a post hoc sensitivity analysis adjusted instead for the use of mechanical ventilation, which has been previously shown to be a valid proxy for stroke severity.[27] In addition, we also examined whether LOS differences by cohorts were associated with inpatient mortality rate differences; for this, we re-estimated the LOS regression model, including inpatient mortality as a covariate. Because of space limitation, additional findings are reported in an Appendix in the online-only Data Supplement. All the statistical analyses were performed using Stata Version 12.1. This study was approved by the Boston University Institutional Review Board.

Results

We identified 147 780 hospitalizations for acute ischemic stroke during 2007 in the 8 states. Their distribution by race and ethnicity was whites=67%, blacks=14%, Hispanics=11%, and others=8% (Table 1).

Patient Risk Factors by Race/Ethnicity and SES

Comparing whites with minorities, we found relatively larger differences in age and prevalence of atrial fibrillation (Table 1). Although 58% of whites were ≥75 years of age, this proportion was 29% and 37% for blacks and Hispanics, respectively ($P<0.001$). Among individuals 18 to 64 years of age, whites comprised 23% of all admissions, compared with 48% and 40% among blacks and Hispanics, respectively. Prevalence of atrial fibrillation was significantly higher among whites (27%) than blacks (12%) and Hispanics (15%; $P<0.001$). This pattern in the 2 factors was persistent in every state (Figure). Patients also exhibited considerable differences by SES (Table 1). Low-income area patients were more often younger and had a higher prevalence of hypertension, diabetes mellitus, and congestive heart failure, but lower prevalence of atrial fibrillation and coronary heart disease.

Hospital Characteristics

Systematic differences were also found in the hospitals where whites and minorities received care (Table 1). Of the 1282 hospitals with acute ischemic stroke admissions in the

Table 1. Patient Risk Factors, Hospital Characteristics, and Patient Outcomes by Race/Ethnicity and Socioeconomic Status, 2007

		Race/Ethnicity			Socioeconomic Status		
	All	Whites	Blacks	Hispanics	Low Income	Medium Income	High Income
No. Admissions	147 780	98 512	20 579	16 927	40 233	38 579	68 968
Patient risk factors							
Female %	53.7	54.1	56.4	50.2	53.7	53.9	53.4
Age (y), %							
18–64	29.4	23.1	47.5	39.9	34.8	28.8	26.4
65–74	20.7	19.4	23.8	22.9	21.9	20.7	20.1
75–84	29.8	33.0	19.4	25.0	26.8	30.3	31.2
85+	20.2	24.5	9.2	12.2	16.5	20.1	22.3
Comorbidity, prevalence, %							
Atrial fibrillation	23.2	27.2	12.3	15.4	19.7	22.9	25.3
Hypertension	77.6	75.0	85.6	80.8	79.2	77.5	76.7
Diabetes mellitus	34.2	28.9	44.4	49.2	39.0	34.5	31.2
Coronary heart disease	27.5	29.7	21.6	24.4	26.8	28.2	27.5
Congestive heart failure	13.7	14.2	14.8	11.8	14.1	13.8	13.5
Charlson score							
0	47.8	47.7	46.2	48.7	46.6	47.5	48.7
1	14.6	15.7	12.8	12.8	14.6	15.0	14.4
2	24.1	23.0	26.9	26.3	25.0	24.0	23.7
3+	13.4	13.6	14.1	12.1	13.8	13.6	13.1
Socioeconomic status (patient residence zip code), %							
Low income area	27.2	19.6	50.8	43.7			
Medium income area	26.1	27.5	21.5	24.5			
High income area	46.7	52.9	27.7	31.8			
Race and ethnicity, %							
Whites, non-Hispanic					48.0	70.2	75.6
Blacks, non-Hispanic					26.0	11.5	8.3
Hispanic					18.4	10.7	7.8
Hospital characteristics							
No. hospitals	1282	1248	954	892	1076	1160	1089
Bed size (No. inpatient beds), %							
≤199	75.2	78.0	72.6	63.7	23.3	28.8	22.2
200–499	15.0	14.1	13.6	23.1	48.6	50.7	52.9
≥500	9.8	7.9	13.8	13.2	28.1	20.5	24.9
ICU beds, %							
≤14	24.2	26.3	16.5	23.1	29.6	34.0	28.1
15 to 36	51.2	52.0	47.9	50.9	33.1	36.5	40.9
≥37	24.7	21.8	35.5	26.0	37.3	29.6	31.0
Nursing staff (No. patients/nurse), %							
≤2	36.6	35.8	37.5	40.1	34.7	33.4	39.6
2.01–3.00	36.4	35.7	36.4	39.4	37.3	36.8	35.5
3.01– 4.00	18.5	19.3	17.6	15.4	18.8	19.7	17.6
>4.00	8.5	9.2	8.5	5.1	9.2	10.1	7.3
Average daily ED volume (No. admissions), %							
≤80	24.2	25.4	18.8	24.2	25.2	27.9	21.5
81–175	45.9	47.2	40.1	44.4	41.6	44.2	49.4
≥176	29.9	27.4	41.1	31.4	33.1	27.9	29.1
Ownership, %							
Nonprofit	75.2	78.0	72.6	63.7	71.9	73.3	78.1

(continued)

Table 1. Continued

		Race/Ethnicity			Socioeconomic Status		
	All	Whites	Blacks	Hispanics	Low Income	Medium Income	High Income
No. Admissions	147 780	98 512	20 579	16 927	40 233	38 579	68 968
For profit	15.0	14.1	13.6	23.1	16.6	16.1	13.5
Government, nonfederal	9.8	7.9	13.8	13.2	11.5	10.7	8.3
Teaching hospital, %	21.8	18.8	34.9	19.4	26.9	18.8	20.5
Safety-net hospital, %	30.6	24.7	41.4	45.2	44.1	30.9	22.6
Patient outcomes							
Inpatient mortality (%)	5.5	5.8	4.4	4.5	5.5	5.4	5.6
Length of stay							
Days, mean	6.2	5.9	7.4	6.3	6.7	6.0	6.0
Frequency of stays >4 d, %	48.3	46.6	55.2	48.5	51.2	47.4	47.0

All acute ischemic discharges for patients age ≥18 years of age in all nonfederal hospitals in AZ, CA, FL, MA, NJ, NY, PA, and TX. Of the 147 780 discharges, 11 762 were for other race and ethnicity. Owing to heterogeneity of this cohort, characteristics are not reported here. However, the other cohort was included in all the statistical analyses. Charlson score excludes congestive heart failure, cerebrovascular, and diabetes mellitus because these are individually specified. ED indicates emergency department; and ICU, intensive care unit.

P value is for the test of comparison across race and ethnicity.

8 states, all admissions for blacks and Hispanics occurred in 954 and 892 hospitals, respectively. In terms of hospital size, blacks and low-income area patients more frequently received care in teaching and safety-net hospitals and hospitals with higher bed size, more intensive care unit beds and emergency department volume, compared with whites and Hispanics.

Differences in Inpatient Mortality Rate

The overall average crude inpatient mortality rate was 5.5%; it was significantly lower for blacks (odds ratio [OR], 0.75; 95% CI, 0.70–0.80) and Hispanics (OR, 0.76; 95% CI, 0.70–0.82) than for whites (Tables 1 and 2). The hierarchical logistic regression model to adjust for patient risk factors indicated very good discrimination (c-statistic=0.74); it also indicates

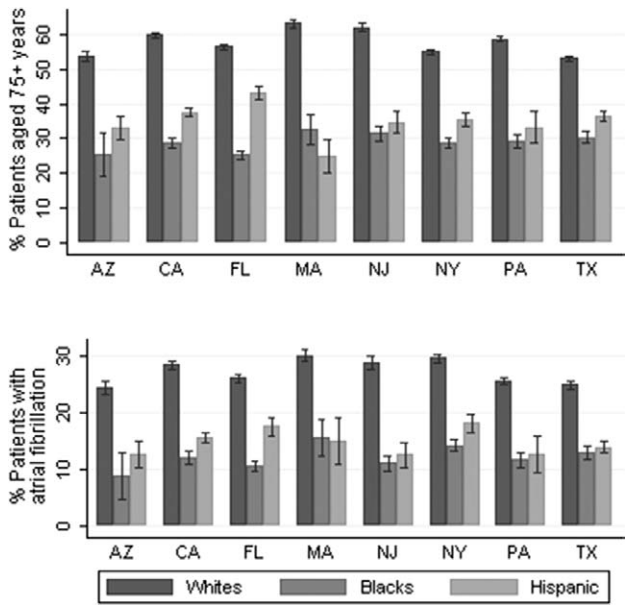


Figure. Prevalence of older age and atrial fibrillation by race and ethnicity, 2007. Note: reported figures are average prevalence rates (%) and 95% confidence intervals.

that only a small share (3.0%) of the residual inpatient mortality differences across patients were associated with the hospital where patients were treated (Appendix in the online-only Data Supplement). Adjusting for patient risk factors, inpatient mortality rates for blacks (OR, 1.02; 95% CI, 0.94–1.10) and Hispanics (OR, 0.98; 95% CI, 0.91–1.07) were similar to that for whites. In assessing the impact of different risk factors in mitigating observed differences in inpatient mortality, we note that whites have higher prevalence of older age, atrial fibrillation, coronary heart disease, and congestive heart failure; in addition, hypertension and diabetes mellitus, conditions for which prevalence was higher among minorities, were either protective or equivocal of inpatient mortality risk.

We also performed comparisons by race/ethnicity separately among patients stratified by SES. Across all 3 SES cohorts, crude inpatient mortality rates were significantly lower for blacks and Hispanics relative to that for whites (Appendix in the online-only Data Supplement). However, adjusted inpatient mortality was generally similar by race/ethnicity among all SES strata; one exception was medium area income Hispanics who had significantly lower mortality (OR, 0.80; 95% CI, 0.67—0.96) relative to whites. We repeated this comparison analysis separately for each of the 8 states and found similarity in risk-adjusted inpatient mortality rates among blacks in all states and among Hispanics in 7 states; however, because of smaller cohort sizes, precision of estimates was lower (Appendix in the online-only Data Supplement).

Similar comparison by SES indicated a converse pattern; although crude inpatient mortality rates were similar by SES cohorts, adjusted rate was higher among low-income area patients (OR, 1.08; 95% CI, 1.02–1.15) when compared with high-income area patients (Table 2). A similar pattern of higher adjusted inpatient mortality for low-income area patients was found among whites, blacks, and Hispanics, although statistical significance was present only for whites, the largest cohort.

The hierarchical logistic regression model to adjust for hospital structural factors and patient risk factors indicated similar discrimination (c-statistic=0.74) compared with the model

Table 2. Unadjusted and Adjusted Inpatient Mortality Rates by Race/Ethnicity and Socioeconomic Status

			Relative Inpatient Mortality Rates								
			Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				95% Confidence Interval			95% Confidence Interval			95% Confidence Interval	
Race/Ethnicity or SES Stratum	N	Observed Inpatient Mortality, %	Odds Ratio	Low	High	Odds Ratio	Low	High	Odds Ratio	Low	High
Race/ethnicity											
White	98 512	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
Black	20 579	4.4	0.75	0.70	0.80	1.02	0.94	1.1	0.96	0.89	1.04
Hispanic	16 927	4.5	0.76	0.70	0.82	0.98	0.91	1.07	0.97	0.89	1.06
Socioeconomic status											
Low income	40 233	5.5	0.99	0.93	1.04	1.08	1.02	1.15	1.06	1.00	1.13
Medium income	38 579	5.4	0.97	0.92	1.02	1.00	0.94	1.06	0.99	0.93	1.06
High income	68 968	5.6	1.00	Reference		1.00	Reference		1.00	Reference	

Adjusted mortality rates obtained by hierarchical logistic regression of inpatient mortality on patient factors (age, sex, and comorbidities) and hospital characteristics (nursing staff, intensive care unit beds, teaching status, and safety-net hospital indicator). In addition, indicators of cohorts compared (race/ethnicity or SES) were included. Separate regression models were estimated for each population stratum examined. Indicator for state was included in all regressions. SES indicates socioeconomic status.

adjusting for patient factors. Further, additional adjustment for hospital factors had little impact on the relative differences in inpatient mortality compared with those obtained by adjusting only for patient risk factors.

Differences in LOS

Although overall mean LOS was 6.2 days, differences by race/ethnicity were significant: whites=5.9, blacks=7.4, and Hispanics=6.3 (P=0.003; Table 1). Correspondingly, rate of longer LOS (LOS >4 days, the median LOS) was significantly higher among blacks (OR, 1.41; 95% CI, 1.37–1.46) and Hispanics (OR, 1.08; 95% CI, 1.04–1.11) compared with whites (Table 3). The regression model to adjust for patient risk factor indicated good discrimination (c-statistic=0.70; Appendix in the online-only Data Supplement). Adjusting for patient risk factors, the relative rate of longer LOS remained higher among blacks (OR, 1.36; 95% CI, 1.31–1.41) and Hispanics (OR, 1.19; 95% CI, 1.14–1.24). This pattern of higher rate of longer stays among minorities was persistent among subpopulations stratified by SES. Further, comparison among all patients of rates of longer LOS indicated higher crude and adjusted rates of longer stay among low SES patients (OR, 1.15; 95% CI, 1.12–1.19) than among high SES patients; this pattern was persistent separately among whites, blacks, and Hispanics (Appendix in the online-only Data Supplement). Further adjustment to also account for differences in hospital characteristics, using hierarchical logistic regression model, indicated little change in the relative rates of longer stay by race/ethnicity and SES. Sensitivity analyses indicated that aforementioned cohort differences in LOS were not associated with differences in inpatient mortality.

Table 3. Unadjusted and Adjusted Rate of Longer Length of Stay (>4 days) by Race/Ethnicity and Socioeconomic Status

			Relative Rates of Longer Stays								
			Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				95% Confidence Interval			95% Confidence Interval			95% Confidence Interval	
Race/Ethnicity or Socioeconomic Stratum	N	Observed Rate of Longer Stays, %	Odds Ratio	Low	High	Odds Ratio	Low	High	Odds Ratio	Low	High
Race/ethnicity											
White	98512	46.6	1.00	Reference		1.00	Reference		1.00	Reference	
Black	20 579	55.2	1.41	1.37	1.46	1.36	1.31	1.41	1.35	1.30	1.40
Hispanic	16 927	48.5	1.08	1.04	1.11	1.19	1.14	1.24	1.18	1.13	1.23
Socioeconomic status											
Low income	40 233	51.2	1.18	1.16	1.21	1.15	1.12	1.19	1.15	1.11	1.19
Medium income	38 579	47.4	1.02	0.99	1.04	1.05	1.02	1.08	1.05	1.02	1.09
High income	68 968	47.0	1.00	Reference		1.00	Reference		1.00	Reference	

Adjusted mortality rates obtained by hierarchical logistic regression of inpatient mortality on patient factors (age, sex, and comorbidities) and hospital characteristics (nursing staff, intensive care unit beds, ownership type, teaching status, and safety net status). Indicator for state was included in both regressions.

Role of Older Age and Atrial Fibrillation

In mitigating the observed differences in inpatient mortality, we evaluated the relative role of the 2 risk factors, older age and presence of atrial fibrillation, whose prevalence was markedly lower among blacks and Hispanics compared with whites (Table 4). Re-estimating the regression, with adjustment made only for these 2 patient factors, we found similar inpatient mortality by race/ethnicity. In contrast, a converse specification with adjustment made for all patient factors excluding age and presence of atrial fibrillation indicated persistence of lower inpatient mortality among blacks (OR, 0.76; 95% CI, 0.71–0.83) and Hispanics (OR, 0.83; 95% CI, 0.78–0.91; Table 4).

Sensitivity Analyses

When added to the model, the use of mechanical ventilation was highly correlated with inpatient death (OR, 28.9; 95% CI, 27.1–30.9; Appendix in the online-only Data Supplement). The pattern of disparities in adjusted inpatient mortality rates was different from that without including the use of mechanical ventilation; adjusted inpatient mortality rates are significantly lower for blacks and Hispanics (compared with whites) but similar for lower income area patients (compared with high-income area patients).

Discussion

Comparing patient outcomes from acute ischemic stroke by race/ethnicity and SES, this study indicates that after adjusting for patient risk factors: (1) inpatient mortality rates are similar among whites, blacks, and Hispanics, but 8% higher among low-income area patients compared with high-income area patients and (2) rates of LOS >4 days are higher among blacks (+36%) and Hispanics (+19%) compared with whites, and among low-income area patients (+15%) compared with high-income area patients. Minorities and lower income area patients more frequently received care at hospitals with larger bed capacity and patient volumes and those with teaching and safety-net status. However, differences in hospital setting were not associated with differences in patient outcomes by race/ethnicity or SES.

Compared with whites, although observed inpatient mortality rates were 27% and 25% lower among blacks and Hispanics, respectively, they were almost completely mitigated after adjusting for patient risk factors. Among the risk factors that differed between whites and minorities, most significant were older age and atrial fibrillation, both of which were markedly less prevalent among blacks and Hispanics. We found that adjusting only for age and atrial fibrillation contributed to most of the mitigation of differences in inpatient mortality by race/ethnicity.

Our findings of lower observed inpatient mortality rates among blacks are qualitatively similar to previously reported differences in stroke subtypes among blacks and whites.[14,28] One recent study, based on brain imaging, found fewer cardioembolic strokes and fewer strokes from large vessel atherosclerosis among blacks compared with whites.[28] As in our study, blacks were younger and had a higher prevalence of hypertension and diabetes mellitus and a lower prevalence of atrial fibrillation, which lends support to the etiologic subtype of small vessel disease.[5,14,29]

Comparisons performed for subpopulations stratified by SES produced the same pattern of racial and ethnic differences in observed inpatient mortality rates, which were nearly completely mitigated after adjustment for patient risk factors (except for medium-income area Hispanics). Our findings of higher rates of LOS >4 days among blacks and Hispanics, relative to whites, are consistent with previous findings.[12,13]

In contrasting our findings with those from previous studies, we distinguish studies based on population-representative data sources (for instance, administrative data) from those based on convenience data (for instance, registry of voluntarily participating providers). Evidence from the former data source type indicates mixed patterns of similar[30] or lower[13,30,31] inpatient mortality among minorities. Of particular relevance are studies that examined administrative data from some of the same states as in this study.[13,30] Using administrative data from New York (2005–2006), a recent study found lower risk-adjusted inpatient mortality among blacks (OR, 0.77; 95% CI, 0.61–0.98) compared with whites.[13] This contrasts with the finding of similar mortality among blacks (OR, 1.0; 95%,

Table 4. Relative Contribution of Risk Factors in Model Prediction of Inpatient Mortality and Length of Stay >4 Days

	Odds Ratio (OR)					
Patient Cohort	Model 1 Unadjusted	Model 2 Adjusted for State Residence	Model 3 Adjusted for Patient Age and State Residence	Model 4 Adjusted for Patient Age, Atrial Fibrillation, and State Residence	Model 5 Adjusted for All Patient Risk Factors and State Residence	Model 6 Adjusted for All Risk Factors Excluding Patient Age and Atrial Fibrillation
Inpatient mortality						
White	Reference	Reference	Reference	Reference	Reference	Reference
Black	0.75***	0.74***	0.94	1.02	0.98	0.76***
Hispanic	0.76***	0.74***	0.89**	0.95	0.98	0.83***
Length of stay >4 d						
White	Reference	Reference	Reference	Reference	Reference	Reference
Black	1.41***	1.32***	1.44***	1.52***	1.43***	1.30***
Hispanic	1.08***	1.18***	1.25***	1.31***	1.29***	1.21***

Models 2 and 4 are based on logistic regressions and models 5 and 6 are based on hierarchical logistic regressions (with hospital level clustering).
***P<0.001; **P<0.01; and *P<0.05.

0.87–1.17) for the New York patient subgroup in our study; this may be because of differences in study design and methods. Specifically, that study excluded 39% of admissions from 57% of hospitals based on rural hospital location or low hospital volume (<10) of black and white stroke patients; in contrast, no such exclusions are made in the current study because our interest is in epidemiological outcome differences regardless of the hospital setting where care was received. Evidence using convenience data on disparities in outcomes is also mixed.[12,14,32] On the basis of data from a large national registry of 1181 hospitals, a recent study found lower inpatient mortality among blacks and similar mortality among Hispanics, relative to that for whites.[12]

Our findings of higher risk-adjusted inpatient mortality among lower income area patients are consistent with most previous studies, although these studies are based on non-US data.[33,34] Mechanisms underlying this association are not well known and seem complex, as noted in the broader literature on disparities by income.[35,36] Unobserved patient factors, including higher stroke severity or lower health literacy, and different attitudes toward life-sustaining therapies among lower income area populations, may underlie higher risk-adjusted inpatient mortality; however, this finding has also been noted in other studies using more detailed clinical indications, including stroke severity.[37]

Our finding of more frequent LOS >4 days among blacks and Hispanics, compared with whites is consistent with previous studies.[12,13] Stratified analysis also indicated that these differences are prevalent among all 3 income area groups. The factors underlying these differences are now well understood. Longer LOS among minorities could be attributable to unmeasured confounders of small vessel stroke or socioeconomic factors (including insurance coverage) that may affect timing of discharge to home or subacute care.[38]

Previous studies have noted the potential role of hospital factors in modifying subgroup differences in stroke outcomes.[12,13,34] Racial and ethnic differences in structural characteristics of hospitals found in this study mirror those in the other studies: minority patients are more likely to receive care in larger hospitals (number of beds, number of intensive care unit beds, and admission volume), including teaching and safety-net hospitals.[12,13] However, these differences were not associated with a measurable change in the pattern of differences in outcomes by race/ethnicity or SES.

This study has several limitations. First, the administrative data used contain limited clinical information on stroke severity. Sensitivity analysis with use of mechanical ventilation as a proxy indicator of patient severity affects the trends in disparities for adjusted inpatient mortality, indicating lower rates for blacks and Hispanics compared with whites and similar rates by SES. The use of mechanical ventilation as measured in the study may not be an accurate proxy of initial stroke severity because it is not possible to distinguish between mechanical ventilation on admission versus a later complication. When present at admission, it may be a valid marker of severe stroke or perhaps of severe comorbid lung disease which is decompensated even by minor stroke. Alternatively, if it results from poor oral hygiene, inadequate dysphagia screening and aspiration pneumonia, then adjusting for its presence may in fact adjust away markers of poor stroke care. Nevertheless, these results point to the potential sensitivity of our main findings to unobserved stroke severity. Another limitation is that the diagnosis codes are not confirmed by patient charts.[39] However, we found that patient differences in only 2 of the risk factors, age and presence of atrial fibrillation, mitigated most of the observed differences in inpatient mortality; additional patient comorbidity indicators did not influence relative differences markedly. Also, our data do not capture important behavioral risk factors (smoking, physical activity, and obesity).[2] Not only is smoking strongly associated with stroke severity, but also its prevalence varies considerably by race/ethnicity and SES. Our cases of acute ischemic strokes, identified by administrative diagnosis codes, were not confirmed by patient charts.[40] Another limitation is that our data cannot distinguish first strokes from secondary strokes. Even though secondary strokes are associated with higher mortality,[21] previous evidence indicates only a moderate black–white difference (33% vs 30%) in the rate of secondary strokes.[12] Also, our data do not include out-of-hospital strokes. Two previous studies reported similar rates (nearly 10%) of strokes (all subtypes) ascertained only in out-of-hospital settings[6,28]; these rates were similar among whites and blacks.[28] The accuracy and consistency of race and ethnicity information, including the extent to which they are patient reported, are also likely to vary across hospitals. Our selection of the 8 states was partly based on completeness of race and ethnicity data; also, all these states had adopted the current federal guidelines for reporting race and ethnicity information.[17,23] In addition, our measure of SES is not an individual measure but based on area-level (zip code) income; also, zip codes are more heterogeneous than census tract or block area units,[41] the latter are not available in the inpatient discharge data.

To summarize, we found that the significantly lower rates of observed inpatient mortality among blacks and Hispanics, compared with that among whites, are primarily associated with differences in patient risk factors; specifically, adjusting for younger age and lower rate of atrial fibrillation among blacks and Hispanics completely mitigates the differences in observed inpatient mortality. Also, we found significantly higher risk-adjusted rates among low-income area patients compared with high-income area patients. However, this pattern was not evident in sensitivity analyses, including the use of mechanical ventilation as a partial surrogate for stroke severity. Further research in potential differences in patient severity or process of care is needed to identify the sources of this differential outcome rate by race/ethnicity and SES.

Sources of Funding

This research has been funded by National Institutes of Health grants 1R21NS062677 (PI=Dr Hanchate) and 1R01HL106029-01 (principal investigator=Dr Hylek).

Disclosures

Dr Schwamm is Chair of the Get With The Guidelines Steering Committee (unpaid) and a consultant to the Massachusetts Department of Public Health. The views expressed in this article are those of the authors and do not necessarily represent the views of the National Institutes of Health, Boston University, or Department of Veterans Affairs. The other authors have no conflicts to report.

References

1. NHLBI. *Morbidity & Mortality: 2009 Chart Book on Cardiovascular, Lung, and Blood Diseases*. Bethesda, MD: National Institutes of Health; 2009.
2. Roger VL, Go AS, Lloyd-Jones DM, Benjamin EJ, Berry JD, Borden WB, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics–2012 update: a report from the American Heart Association. *Circulation*. 2012;125:e2–e220.
3. National Center for Health Statistics. *Health, United States, 2010: With Special Feature on Death and Dying*. Hyattsville, Maryland: Prevention City; 2011.
4. Broderick JP, Ferioli S. Assessing stroke incidence and mortality across the United States. *Ann Neurol*. 2011;69:595–596.
5. Howard VJ, Kleindorfer DO, Judd SE, McClure LA, Safford MM, Rhodes JD, et al. Disparities in stroke incidence contributing to disparities in stroke mortality. *Ann Neurol*. 2011;69:619–627.
6. Kissela B, Schneider A, Kleindorfer D, Khoury J, Miller R, Alwell K, et al. Stroke in a biracial population: the excess burden of stroke among blacks. *Stroke*. 2004;35:426–431.
7. Kleindorfer D, Broderick J, Khoury J, Flaherty M, Woo D, Alwell K, et al. The unchanging incidence and case-fatality of stroke in the 1990s: a population-based study. *Stroke*. 2006;37:2473–2478.
8. Rosamond WD, Folsom AR, Chambless LE, Wang CH, McGovern PG, Howard G, et al. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. *Stroke*. 1999;30:736–743.
9. Morgenstern LB, Smith MA, Lisabeth LD, Risser JM, Uchino K, Garcia N, et al. Excess stroke in Mexican Americans compared with non-Hispanic Whites: the Brain Attack Surveillance in Corpus Christi Project. *Am J Epidemiol*. 2004;160:376–383.
10. White H, Boden-Albala B, Wang C, Elkind MS, Rundek T, Wright CB, et al. Ischemic stroke subtype incidence among whites, blacks, and Hispanics: the Northern Manhattan Study. *Circulation*. 2005;111:1327–1331.
11. Sacco RL, Boden-Albala B, Gan R, Chen X, Kargman DE, Shea S, et al. Stroke incidence among white, black, and Hispanic residents of an urban community: the Northern Manhattan Stroke Study. *Am J Epidemiol*. 1998;147:259–268.
12. Schwamm LH, Reeves MJ, Pan W, Smith EE, Frankel MR, Olson D, et al. Race/ethnicity, quality of care, and outcomes in ischemic stroke. *Circulation*. 2010;121:1492–1501.
13. Xian Y, Holloway RG, Noyes K, Shah MN, Friedman B. Racial differences in mortality among patients with acute ischemic stroke: an observational study. *Ann Intern Med*. 2011;154:152–159.
14. Stansbury JP, Jia H, Williams LS, Vogel WB, Duncan PW. Ethnic disparities in stroke: epidemiology, acute care, and postacute outcomes. *Stroke*. 2005;36:374–386.
15. Kuhlemeier KV, Stiens SA. Racial disparities in severity of cerebrovascular events. *Stroke*. 1994;25:2126–2131.
16. Lisabeth LD, Risser JM, Brown DL, Al-Senani F, Uchino K, Smith MA, et al. Stroke burden in Mexican Americans: the impact of mortality following stroke. *Ann Epidemiol*. 2006;16:33–40.
17. Schoenman JA, Sutton JP, Kintala S, Love D, Maw R. *The Value of Hospital Discharge Databases.* Chicago, IL: NORC at the University of Chicago; 2005.
18. U.S. Census Bureau. Population Profile of the United States: 2000 (Internet Release). Washington, DC: U.S. Census Bureau; 2008.
19. American Hospital Association. *AHA Annual Survey Database*. Chicago, IL: American Hospital Association; 2010.
20. Agency for Healthcare Research and Quality. *Guide to Inpatient Quality Indicators: Quality of Care in Hospitals - Volume, Mortality and Utilization*. Agency for Healthcare Research and Quality; 2011.
21. Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics–2011 update: a report from the American Heart Association. *Circulation*. 2011;123:e18–e209.
22. Goldstein LB, Samsa GP, Matchar DB, Horner RD. Charlson Index comorbidity adjustment for ischemic stroke outcome studies. *Stroke*. 2004;35:1941–1945.
23. Federal Registrar. Revisions to the standards for the classification of federal data on race and ethnicity. *Federal Registrar*. 1997;62:58781–58790
24. Jha AK, Fisher ES, Li Z, Orav EJ, Epstein AM. Racial trends in the use of major procedures among the elderly. *N Engl J Med*. 2005;353:683–691.
25. Skinner J, Weinstein JN, Sporer SM, Wennberg JE. Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients. *N Engl J Med*. 2003;349:1350–1359.
26. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? *J Am Coll Surg*. 2004;198:626–632.
27. Horner RD, Sloane RJ, Kahn KL. Is use of mechanical ventilation a reasonable proxy indicator for coma among Medicare patients hospitalized for acute stroke? *Health Serv Res*. 1998;32:841–859.
28. Markus HS, Khan U, Birns J, Evans A, Kalra L, Rudd AG, et al. Differences in stroke subtypes between black and white patients with stroke: the South London Ethnicity and Stroke Study. *Circulation*. 2007;116:2157–2164.
29. Hajat C, Tilling K, Stewart JA, Lemic-Stojcevic N, Wolfe CD. Ethnic differences in risk factors for ischemic stroke: a European case-control study. *Stroke*. 2004;35:1562–1567.
30. Polsky D, Lave J, Klusaritz H, Jha A, Pauly MV, Cen L, et al. Is lower 30-day mortality posthospital admission among blacks unique to the Veterans Affairs health care system? *Med Care*. 2007;45:1083–1089.
31. Volpp KG, Stone R, Lave JR, Jha AK, Pauly M, Klusaritz H, et al. Is thirty-day hospital mortality really lower for black veterans compared with white veterans? *Health Serv Res*. 2007;42:1613–1631.
32. Wolfe CD, Smeeton NC, Coshall C, Tilling K, Rudd AG. Survival differences after stroke in a multiethnic population: follow-up study with the South London stroke register. *BMJ*. 2005;331:431.
33. Cox AM, McKevitt C, Rudd AG, Wolfe CD. Socioeconomic status and stroke. *Lancet Neurol*. 2006;5:181–188.
34. Saposnik G, Jeerakathil T, Selchen D, Baibergenova A, Hachinski V, Kapral MK; Stroke Outcome Research Canada Working Group. Socioeconomic status, hospital volume, and stroke fatality in Canada. *Stroke*. 2008;39:3360–3366.
35. Fiscella K, Franks P, Gold MR, Clancy CM. Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. *J Am Med Assoc*. 2000;283:2579–2584.
36. Gornick ME, Eggers PW, Reilly TW, Mentnech RM, Fitterman LK, Kucken LE, et al. Effects of race and income on mortality and use of services among Medicare beneficiaries. *N Engl J Med*. 1996;335:791–799.
37. Borrell LN, Diez Roux AV, Rose K, Catellier D, Clark BL; Atherosclerosis Risk in Communities Study. Neighbourhood characteristics and mortality in the Atherosclerosis Risk in Communities Study. *Int J Epidemiol*. 2004;33:398–407.
38. van Straten A, van der Meulen JH, van den Bos GA, Limburg M. Length of hospital stay and discharge delays in stroke patients. *Stroke*. 1997;28:137–140.
39. Schoenman JA, Sutton JP, Elixhauser A, Love D. Understanding and enhancing the value of hospital discharge data. *Med Care Res Rev*. 2007;64:449–468.
40. Goldstein LB. Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke: effect of modifier codes. *Stroke*. 1998;29:1602–1604.
41. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health*. 1997;18:341–378.

SUPPLEMENTAL MATERIAL

Appendix

Comparison of Ischemic Stroke Outcomes and, Patient and Hospital Characteristics by Race/Ethnicity and Socioeconomic Status

Table A1. Comparison Included and Excluded Observations

(N=155,539 discharges for Acute Ischemic Stroke in all non-federal hospitals in AZ, CA, FL, MA, NJ, NY, PA and TX, 2007)

	Included	Excluded
# admissions	147,780	7759
Female%	53.7	42.8
Age (%)		
18-64	29.4	41.2
65-74	20.7	20.5
75-84	29.8	23.9
85+	20.2	14.4
Race/ethnicity (%)		
Whites	66.7	62.9
Blacks	13.9	17.4
Hispanics	11.5	14.1
Other	8.0	5.6
Comorbidity (%)		
Atrial fibrillation	23.2	18.1
Hypertension	77.6	75.9
Diabetes	34.2	32.3
Coronary Heart Disease	27.5	23.8
Congestive Heart Failure	13.7	12.3
Charlson Score		
0	47.8	47.5
1	14.6	13.8
2	24.1	25.0
3+	13.4	13.7
Inpatient Mortality (%)	5.5	4.6
LOS, Median	4	4

Table A2. Inpatient Mortality Risk Factors: Regression Estimates

	Model 1 Adjusted for Patient Characteristics			Model 2 Adjusted for Patient & Hospital Characteristics		
	Odds Ratio	95% CI Low	95% CI High	Odds Ratio	95% CI Low	95% CI High
Age: 18-64	1.0	Reference		1.0	Reference	
65-74	1.26***	1.17	1.37	1.27***	1.18	1.38
75-84	1.54***	1.44	1.66	1.56***	1.45	1.68
85+	2.18***	2.02	2.35	2.21***	2.05	2.38
Female	1.02	0.97	1.07	1.03	0.98	1.08
Comorbidity						
Diabetes	0.97	0.93	1.03	0.97	0.92	1.03
Hypertension	0.63***	0.6	0.66	0.63***	0.6	0.66
Congestive Heart Failure	1.80***	1.7	1.9	1.79***	1.7	1.9
Coronary Heart Disease	1.17***	1.11	1.23	1.17***	1.11	1.23
Atrial fibrillation	1.81***	1.72	1.9	1.81***	1.72	1.9
Charlson Score						
0	1.0	Reference		1.0	Reference	
1	1.21***	1.13	1.31	1.22***	1.13	1.31
2	1.48***	1.39	1.57	1.47***	1.39	1.56
3+	2.15***	2.02	2.29	2.15***	2.02	2.3
State						
Arizona	0.43***	0.35	0.53	0.43***	0.34	0.54
California	1.0	Reference		1.0	Reference	
Florida	0.66***	0.58	0.74	0.68***	0.6	0.77
Massachusetts	1.07	0.91	1.26	1.07	0.9	1.27
New Jersey	0.88	0.75	1.02	0.9	0.77	1.05
New York	1.29***	1.15	1.44	1.24***	1.1	1.4
Pennsylvania	0.75***	0.66	0.85	0.78***	0.68	0.88
Texas	0.84**	0.75	0.94	0.85**	0.76	0.96
ICU Beds						
<= 14				1.0	Reference	
15 to 36				0.93	0.86	1.01
>= 37				0.95	0.85	1.06
Nursing staff (# patients/nurse)						
<= 2				1.0	Reference	
2.01 to 3.00				0.96	0.87	1.04
3.01 to 4.00				0.95	0.85	1.06
> 4.00				0.99	0.86	1.13
Ownership						
Non Profit				1.0	Reference	
For Profit				0.91	0.82	1.02
Government, non-Federal				1.20**	1.07	1.36
Teaching hospital, %				1.12	0.99	1.26
Safety Net hospital, %				1.14**	1.05	1.24
Constant	0.03***	0.03	0.04	0.03***	0.03	0.03
C-statistic	0.74			0.74		

Note: *, ** and *** denote $p<0.05$, $p<0.01$ and $p<0.001$ respectively

Table A3. Risk Factors for Length of Stay > 4 days: Regression Estimates

	Model 1 Adjust for Patient Characteristics			Model 2 Adjust for Patient & Hospital Characteristics		
	Odds Ratio	95% CI Low	High	Odds Ratio	95% CI Low	High
Age						
18-64	Reference			Reference		
65-74	0.95**	0.92	0.99	0.96**	0.93	0.99
75-84	1.07***	1.04	1.1	1.07***	1.04	1.11
85+	1.04*	1	1.07	1.04*	1	1.08
Female	1.08***	1.05	1.1	1.08***	1.05	1.1
Comorbidity						
Diabetes	1.18***	1.16	1.21	1.18***	1.16	1.21
Hypertension	0.95***	0.92	0.97	0.95***	0.92	0.97
Congestive Heart Failure	1.56***	1.5	1.61	1.56***	1.5	1.61
Coronary Heart Disease	0.99	0.96	1.01	0.98	0.96	1.01
Atrial fibrillation	1.71***	1.66	1.76	1.71***	1.66	1.76
Charlson Score						
0	Reference			Reference		
1	1.40***	1.35	1.45	1.40***	1.35	1.45
2	1.98***	1.93	2.04	1.98***	1.93	2.03
3+	2.66***	2.56	2.75	2.66***	2.56	2.75
State						
Arizona	0.93	0.78	1.1	0.93*	0.87	0.98
California	Reference			Reference		
Florida	1.38***	1.24	1.53	1.39***	1.34	1.44
Massachusetts	0.97	0.83	1.14	1.09**	1.03	1.15
New Jersey	2.26***	1.95	2.62	2.14***	2.04	2.24
New York	2.86***	2.56	3.18	2.66***	2.55	2.76
Pennsylvania	1.47***	1.31	1.65	1.47***	1.41	1.53
Texas	1.32***	1.19	1.46	1.41***	1.36	1.46
Ownership, %						
Non Profit				Reference		
For Profit				1.17***	1.07	1.29
Government, non-Federal				1.03	0.92	1.16
Bed size (# inpatient beds)						
<= 199				Reference		
200 to 499				1.08	0.99	1.18
>= 500				1.26**	1.07	1.48
ICU Beds						
<= 14				Reference		
15 to 36				1.06	0.97	1.16
>= 37				1.15*	1.01	1.31
Nursing staff (# patients/nurse)						
<= 2				Reference		
2.01 to 3.00				1.07	0.99	1.16
3.01 to 4.00				1.16**	1.05	1.28

> 4.00				1.22**	1.08	1.39
Emergency department admissions/day						
<= 80				Reference		
81 to 175				1	0.92	1.09
>= 176				0.95	0.84	1.08
Teaching hospital, %				1.1	0.97	1.26
Safety Net hospital, %				1.01	0.94	1.1
Constant	0.33***	0.31	0.36	0.25***	0.21	0.3

Note: *, ** and *** denote $p<0.05$, $p<0.01$ and $p<0.001$ respectively

Table A4. Unadjusted and Adjusted Inpatient Mortality Rates by Race, Ethnicity and Socioeconomic Status

Population Cohort	Race/Ethnicity or SES Stratum	N	Observed Inpatient Mortality, %	Relative Inpatient Mortality Rates								
				Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				Odds Ratio	95% CI Low	95% CI High	Odds Ratio	95% CI Low	95% CI High	Odds Ratio	95% CI Low	95% CI High
Comparisons by Race & Ethnicity												
All	White	98,512	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	20,579	4.4	0.75	0.70	0.80	1.02	0.94	1.1	0.96	0.89	1.04
	Hisp	16,927	4.5	0.76	0.70	0.82	0.98	0.91	1.07	0.97	0.89	1.06
Low Income	White	19,297	6.0	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	10,455	4.8	0.78	0.70	0.87	1.04	0.92	1.18	1.03	0.90	1.17
	Hisp	7,394	4.9	0.80	0.71	0.91	1.05	0.91	1.2	1.04	0.91	1.20
Medium Income	White	27,063	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	4,427	4.0	0.67	0.57	0.79	0.88	0.74	1.05	0.87	0.73	1.03
	Hisp	4,142	3.8	0.64	0.54	0.76	0.8	0.67	0.96	0.80	0.67	0.96
High Income	White	52,152	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	5,697	4.2	0.71	0.62	0.81	0.95	0.82	1.1	0.93	0.81	1.08
	Hisp	5,391	4.4	0.76	0.66	0.87	0.99	0.86	1.15	0.99	0.85	1.14
Comparisons by Socioeconomic Status												
All	Low Income	40,233	5.5	0.99	0.93	1.04	1.08	1.02	1.15	1.06	1.00	1.13
	Med Income	38,579	5.4	0.97	0.92	1.02	1.00	0.94	1.06	0.99	0.93	1.06
	High Income	68,968	5.6	1.00	Reference		1.00	Reference		1.00	Reference	
Whites	Low Income	19,297	6.0	1.04	0.97	1.11	1.07	0.99	1.16	1.05	0.97	1.14
	Med Income	27,063	5.8	1.01	0.95	1.08	1.05	0.98	1.12	1.04	0.97	1.12
	High Income	52,152	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
Blacks	Low Income	10,455	4.8	1.15	0.98	1.35	1.17	0.99	1.38	1.15	0.97	1.36
	Med Income	4,427	4.0	0.96	0.79	1.17	0.97	0.79	1.19	0.96	0.78	1.18
	High Income	5,697	4.2	1.00	Reference		1.00	Reference		1.00	Reference	
Hispanics	Low Income	7,394	4.9	1.10	0.93	1.30	1.12	0.93	1.34	1.11	0.92	1.33
	Med Income	4,142	3.8	0.86	0.70	1.06	0.85	0.69	1.05	0.84	0.68	1.04
	High Income	5,391	4.4	1.00	Reference		1.00	Reference		1.00	Reference	

Notes: Adjusted mortality rates obtained by hierarchical logistic regression of inpatient mortality on patient factors (age, sex and comorbidities) and hospital characteristics (nursing staff, ICU beds, teaching status and safety net hospital indicator) – as in Table A2

above. In addition indicators of cohorts compared (race/ethnicity or SES) were included. Separate regression models were estimated for each population stratum examined. Indicator for state was included in all regressions.

Table A5. Unadjusted and Adjusted Rate of Longer (>4 days) Length of Stay by Race/Ethnicity and Socioeconomic Status

Population Cohort	Race/Ethnicity or SES Stratum	N	Observed Rate of Longer Stays, %	Relative Rates of Longer Stays								
				Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				Odds Ratio	95% CI Low	High	Odds Ratio	Odds Ratio Low	High	95% CI	Odds Ratio Low	High
Comparisons by Race & Ethnicity												
All	White	98,512	46.6	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	20,579	55.2	1.41	1.37	1.46	1.36	1.31	1.41	1.35	1.30	1.40
	Hisp	16,927	48.5	1.08	1.04	1.11	1.19	1.14	1.24	1.18	1.13	1.23
Low Income	White	19,297	48.1	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	10,455	56.1	1.38	1.31	1.45	1.33	1.25	1.42	1.32	1.24	1.40
	Hisp	7,394	52.0	1.17	1.11	1.23	1.24	1.16	1.33	1.23	1.14	1.32
Medium Income	White	27,063	46.2	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	4,427	54.1	1.37	1.29	1.46	1.37	1.27	1.47	1.35	1.25	1.45
	Hisp	4,142	45.1	0.96	0.90	1.02	1.12	1.03	1.21	1.11	1.03	1.20
High Income	White	52,152	46.2	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	5,697	54.4	1.39	1.31	1.46	1.35	1.27	1.44	1.34	1.26	1.43
	Hisp	5,391	46.1	1.00	0.94	1.05	1.16	1.09	1.24	1.16	1.08	1.24
Comparisons by Socioeconomic Status												
All	Low Income	40,233	51.2	1.18	1.16	1.21	1.15	1.12	1.19	1.15	1.11	1.19
	Med Income	38,579	47.4	1.02	0.99	1.04	1.05	1.02	1.08	1.05	1.02	1.09
	High Income	68,968	47.0	1.00	Reference		1.00	Reference		1.00	Reference	
Whites	Low Income	19,297	48.1	1.08	1.04	1.11	1.07	1.03	1.12	1.08	1.03	1.12
	Med Income	27,063	46.2	1.00	0.97	1.03	1.02	0.98	1.06	1.02	0.98	1.06
	High Income	52,152	46.2	1.00	Reference		1.00	Reference		1.00	Reference	
Blacks	Low Income	10,455	56.1	1.07	1.00	1.14	1.05	0.97	1.14	1.05	0.97	1.13
	Med Income	4,427	54.1	0.99	0.91	1.07	1.00	0.91	1.10	1.00	0.91	1.10
	High Income	5,697	54.4	1.00	Reference		1.00	Reference		1.00	Reference	
Hispanics	Low Income	7,394	52.0	1.27	1.18	1.36	1.18	1.08	1.29	1.16	1.06	1.27
	Med Income	4,142	45.1	0.96	0.89	1.04	1.02	0.93	1.12	1.01	0.92	1.11
	High Income	5,391	46.1	1.00	Reference		1.00	Reference		1.00	Reference	

Notes: Adjusted mortality rates obtained by hierarchical logistic regression of inpatient mortality on patient factors (age, sex and comorbidities) and hospital characteristics (nursing staff, ICU beds, teaching status and safety net hospital indicator) – as in Table A2

above. In addition indicators of cohorts compared (race/ethnicity or SES) were included. Separate regression models were estimated for each population stratum examined. Indicator for state was included in all regressions.

Table A6. Unadjusted and Adjusted Inpatient Mortality Rates By State & Race/Ethnicity

State	Race & Ethnicity	N	Observed Inpatient Mortality, %	Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				Odds Ratio	95% CI Low	95% CI High	Odds Ratio	Odds Ratio Low	Odds Ratio High	95% CI	Odds Ratio Low	Odds Ratio High
AZ	White	4,605	2.6	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	193	2.1	0.78	0.29	2.15	0.85	0.30	2.39	0.79	0.27	2.34
	Hispanic	775	3.1	1.18	0.76	1.85	1.36	0.85	2.20	1.35	0.84	2.16
CA	White	18,615	6.9	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	3,234	4.4	0.62	0.52	0.75	0.84	0.69	1.02	0.82	0.68	0.99
	Hispanic	5,965	4.8	0.68	0.59	0.77	0.92	0.79	1.06	0.89	0.77	1.02
FL	White	17,184	4.3	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	4,217	3.8	0.88	0.74	1.05	1.11	0.91	1.34	1.07	0.88	1.29
	Hispanic	2,387	3.7	0.87	0.69	1.08	0.98	0.76	1.25	1.02	0.81	1.29
MA	White	6,383	7.7	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	471	5.5	0.7	0.47	1.05	1.00	0.65	1.56	0.91	0.59	1.40
	Hispanic	307	3.3	0.4	0.21	0.76	0.66	0.34	1.28	0.67	0.35	1.28
NJ	White	7,041	6.2	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	1,882	3.6	0.56	0.43	0.72	0.79	0.59	1.06	0.72	0.53	0.97
	Hispanic	831	3.6	0.56	0.39	0.82	0.74	0.49	1.10	0.65	0.43	0.98
NY	White	16,265	7.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	4,592	5.6	0.7	0.61	0.8	0.95	0.80	1.11	0.98	0.84	1.14
	Hispanic	2,172	4.9	0.6	0.49	0.74	0.80	0.64	1.01	0.8	0.64	0.99
PA	White	16,079	5.0	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	2,523	3.8	0.76	0.61	0.94	1.10	0.86	1.40	0.98	0.77	1.25
	Hispanic	387	4.4	0.88	0.54	1.44	1.33	0.79	2.23	1.17	0.71	1.92
TX	White	12,340	4.9	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	3,467	4.5	0.92	0.77	1.10	1.09	0.90	1.32	1.1	0.91	1.33
	Hispanic	4,103	4.8	0.98	0.83	1.16	1.25	1.03	1.50	1.25	1.04	1.50

Notes: 1) Adjusted mortality rates obtained by hierarchical logistic regression of inpatient mortality on patient factors (age, sex and comorbidities) and hospital characteristics (nursing staff, ICU beds, teaching status and safety net hospital indicator). Indicator for state was included in both regressions.

Table A7. Unadjusted and Adjusted Inpatient Mortality Rates By State & Socioeconomic Status

State	Income	N	Observed Inpatient Mortality, %	Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				Odds Ratio	95% CI		Odds Ratio	Odds Ratio		95% CI	Odds Ratio	
					Low	High		Low	High		Low	High
AZ	Low	1,360	2.8	1.11	0.75	1.64	1.25	1.07	1.46	1.00	0.65	1.55
	Medium	1,487	3.2	1.28	0.89	1.85	0.99	0.85	1.14	1.07	0.71	1.61
	High	3,040	2.5	1.00	Reference		1.00	Reference		1.00	Reference	
CA	Low	8,492	5.8	0.89	0.79	0.99	1.12	1.04	1.21	1.01	0.90	1.14
	Medium	8,086	5.8	0.89	0.80	1.00	1.09	1.02	1.17	0.98	0.87	1.10
	High	15,449	6.5	1.00	Reference		1.00	Reference		1.00	Reference	
FL	Low	7,044	4.4	1.10	0.95	1.28	1.19	1.1	1.28	1.10	0.94	1.29
	Medium	6,882	4.1	1.02	0.87	1.19	1.04	0.97	1.12	1.03	0.88	1.21
	High	10,721	4.0	1.00	Reference		1.00	Reference		1.00	Reference	
MA	Low	2,053	7.2	0.89	0.72	1.09	1.08	0.94	1.23	0.95	0.76	1.20
	Medium	1,978	6.5	0.79	0.64	0.98	1.04	0.92	1.18	0.80	0.63	1.00
	High	3,524	8.1	1.00	Reference		1.00	Reference		1.00	Reference	
NJ	Low	3,126	4.8	0.81	0.66	0.99	1.32	1.17	1.5	0.88	0.69	1.13
	Medium	2,798	5.6	0.94	0.77	1.15	1.02	0.91	1.14	1.01	0.81	1.26
	High	4,485	5.9	1.00	Reference		1.00	Reference		1.00	Reference	
NY	Low	6,909	6.6	0.87	0.78	0.98	1.08	1	1.17	0.94	0.83	1.07
	Medium	6,623	7.1	0.94	0.84	1.06	0.99	0.92	1.08	0.93	0.82	1.05
	High	12,809	7.5	1.00	Reference		1.00	Reference		1.00	Reference	
PA	Low	5,518	5.2	1.13	0.97	1.32	1.15	1.06	1.26	1.29	1.09	1.52
	Medium	5,105	5.0	1.09	0.93	1.28	1.08	0.99	1.18	1.18	1.00	1.39
	High	9,240	4.6	1.00	Reference		1.00	Reference		1.00	Reference	
TX	Low	5,731	5.9	1.43	1.23	1.65	1.21	1.12	1.32	1.42	1.21	1.66
	Medium	5,620	5.0	1.19	1.02	1.39	1.09	1.01	1.18	1.13	0.97	1.33
	High	9,700	4.2	1.00	Reference		1.00	Reference		1.00	Reference	

Note: 1) Adjusted mortality rates obtained by logistic regression of inpatient mortality on patient factors (age, sex and comorbidities) and hospital characteristics (bed size, ICU beds, nursing staff, average daily ED volume, teaching status and safety net hospital indicator). Indicator for state was included in both regressions.

Table A8. Inpatient Mortality Risk Factors: Adding Use of Mechanical Ventilation

	Model 1			Model 2		
	Odds Ratio	95% CI		Odds Ratio	95% CI	
		Low	High		Low	High
Age						
18-64	1.0	Reference		1.0	Reference	
65-74	1.38***	1.27	1.51	1.38***	1.26	1.51
75-84	1.97***	1.82	2.13	1.96***	1.81	2.12
85+	3.40***	3.13	3.69	3.38***	3.11	3.67
Female	1.09**	1.03	1.15	1.09**	1.03	1.15
Use of mechanical ventilation	28.92***	27.07	30.9	29.18***	27.31	31.18
Comorbidity						
Diabetes	0.97	0.92	1.03	0.97	0.92	1.03
Hypertension	0.70***	0.66	0.74	0.70***	0.66	0.74
Congestive Heart Failure	1.55***	1.46	1.65	1.55***	1.46	1.65
Coronary Heart Disease	1.23***	1.16	1.3	1.23***	1.16	1.3
Atrial fibrillation	1.62***	1.54	1.72	1.63***	1.54	1.72
Charlson Score						
0	1.0	Reference		1.0	Reference	
1	1.15***	1.06	1.25	1.15***	1.06	1.24
2	1.35***	1.27	1.44	1.35***	1.27	1.44
3+	1.94***	1.81	2.08	1.95***	1.81	2.09
State						
Arizona	0.41***	0.32	0.53	0.43***	0.33	0.55
California	1.0	Reference		1.0	Reference	
Florida	0.62***	0.55	0.71	0.66***	0.57	0.75
Massachusetts	1.1	0.92	1.32	1.14	0.94	1.38
New Jersey	0.83*	0.7	0.99	0.88	0.74	1.04
New York	1.17*	1.03	1.32	1.15*	1.01	1.31
Pennsylvania	0.82**	0.72	0.95	0.85*	0.74	0.98
Texas	0.89	0.78	1.01	0.93	0.81	1.05
ICU Beds						
<= 14				1.0	Reference	
15 to 36				0.86**	0.78	0.94
>= 37				0.83**	0.73	0.94
Nursing staff (# patients/nurse)						
<= 2				1.0	Reference	
2.01 to 3.00				0.97	0.88	1.07
3.01 to 4.00				0.96	0.85	1.08
> 4.00				1.05	0.9	1.22
Ownership						
Non Profit				1.0	Reference	
For Profit				0.88*	0.78	1
Government, non-Federal				1.20**	1.05	1.38
Teaching hospital, %				0.9	0.78	1.02
Safety Net hospital, %				1.03	0.94	1.13
Constant	0.02***	0.01	0.02	0.02***	0.02	0.02
C-statistic	0.85			0.85		

Note: *, ** and *** denote p<0.05, p<0.01 and p<0.001 respectivel

Table A9. Unadjusted and Adjusted Inpatient Mortality Rates by Race, Ethnicity and Socioeconomic Status: Adding Use of Mechanical Ventilation

Population Cohort	Race/Ethnicity or SES Stratum	N	Observed Inpatient Mortality, %	Relative Inpatient Mortality Rates								
				Unadjusted			Adjusted for Patient Characteristics			Adjusted for Patient and Hospital Characteristics		
				Odds Ratio	95% Conf. Int Low	High	Odds Ratio	95% Conf. Int Low	High	Odds Ratio	95% Conf. Int Low	High
Comparisons by Race & Ethnicity												
All	White	98,512	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	20,579	4.4	0.75	0.70	0.80	0.83	0.76	0.91	0.84	0.77	0.92
	Hisp	16,927	4.5	0.76	0.70	0.82	0.90	0.82	0.99	0.91	0.82	1.00
Low Income	White	19,297	6.0	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	10,455	4.8	0.78	0.70	0.87	0.83	0.72	0.96	0.86	0.75	1.00
	Hisp	7,394	4.9	0.80	0.71	0.91	0.91	0.78	1.06	0.95	0.81	1.12
Medium Income	White	27,063	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	4,427	4.0	0.67	0.57	0.79	0.73	0.6	0.88	0.74	0.61	0.89
	Hisp	4,142	3.8	0.64	0.54	0.76	0.72	0.59	0.89	0.73	0.6	0.89
High Income	White	52,152	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
	Black	5,697	4.2	0.71	0.62	0.81	0.83	0.71	0.98	0.83	0.71	0.98
	Hisp	5,391	4.4	0.76	0.66	0.87	0.93	0.79	1.09	0.93	0.8	1.09
Comparisons by Socioeconomic Status												
All	Low Income	40,233	5.5	0.99	0.93	1.04	1.04	0.97	1.12	1.04	0.97	1.11
	Med Income	38,579	5.4	0.97	0.92	1.02	1.00	0.93	1.07	0.99	0.92	1.06
	High Income	68,968	5.6	1.00	Reference		1.00	Reference		1.00	Reference	
Whites	Low Income	19,297	6.0	1.04	0.97	1.11	1.07	0.98	1.17	1.05	0.96	1.15
	Med Income	27,063	5.8	1.01	0.95	1.08	1.07	0.99	1.15	1.05	0.98	1.14
	High Income	52,152	5.8	1.00	Reference		1.00	Reference		1.00	Reference	
Blacks	Low Income	10,455	4.8	1.15	0.98	1.35	1.09	0.9	1.32	1.11	0.92	1.36
	Med Income	4,427	4.0	0.96	0.79	1.17	0.91	0.72	1.15	0.91	0.72	1.16
	High Income	5,697	4.2	1.00	Reference		1.00	Reference		1.00	Reference	
Hispanics	Low Income	7,394	4.9	1.10	0.93	1.30	1.08	0.88	1.33	1.13	0.92	1.39
	Med Income	4,142	3.8	0.86	0.70	1.06	0.83	0.66	1.06	0.84	0.66	1.06
	High Income	5,391	4.4	1.00	Reference		1.00	Reference		1.00	Reference	

Notes: Adjusted mortality rates obtained by hierarchical logistic regression of inpatient mortality on patient factors (age, sex and comorbidities) and hospital characteristics (nursing staff, ICU beds, teaching status and safety net hospital indicator). In addition indicators of cohorts compared (race/ethnicity or SES) were included. Separate regression models were estimated for each population stratum examined. Indicator for state was included in all regressions.

Clinical Endocrinology (2013) **78**, 629–634 doi: 10.1111/cen.12013

ORIGINAL ARTICLE

Association of sex steroids, gonadotrophins, and their trajectories with clinical cardiovascular disease and all-cause mortality in elderly men from the Framingham Heart Study

Robin Haring*,†, Zhaoyang Teng‡, Vanessa Xanthakis*,‡,§, Andrea Coviello*,¶, Lisa Sullivan‡, Shalender Bhasin¶, Joanne M. Murabito§,**, Henri Wallaschofski† and Ramachandran S. Vasan*,§

*Preventive Medicine & Epidemiology Section, School of Medicine, Boston University, Boston, MA, USA, †Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany, ‡Department of Biostatistics, School of Public Health, Boston University, Boston, §National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, ¶Section of Endocrinology, Diabetes, and Nutrition, School of Medicine, Boston University, Boston, and **Section of General Internal Medicine, School of Medicine, Boston University, Boston, MA, USA

Summary

Background Emerging data from longitudinal studies suggest that low sex steroid concentrations in men are associated with increased cardiovascular risk and mortality. The impact of longitudinal trajectory patterns from serial sex steroid and gonadotrophin measurements on the observed associations is unknown to date.

Methods We prospectively evaluated 254 elderly men (mean age, 75·5 years) of the Framingham Heart Study with up to four serial measurements of serum total testosterone (TT), dehydroepiandrosterone sulphate (DHEAS), follicle-stimulating hormone (FSH), luteinizing hormone (LH) and total estradiol (EST); and constructed age- and multivariable-adjusted Cox proportional hazard regression models relating baseline hormone concentrations and their mean, slope and variation over time (modelled as continuous and categorized into quartiles) to the incidence of clinical cardiovascular disease (CVD) and all-cause mortality at 5- and 10-year follow-up.

Results We observed no association between baseline concentrations of sex steroids, gonadotrophins and their trajectories with incident clinical CVD over 5- and 10-year follow-up. Although higher baseline TT concentrations were associated with lower mortality risk at 5 years (hazard ratio per quartile increment, 0·74; 95% confidence interval, 0·56–0·98), correction for multiple statistical testing ($P < 0·005$) rendered this association statistically nonsignificant. Repeat analyses at the 10-year follow-up time point also demonstrated no significant association between sex steroids, gonadotrophins or their trajectories and mortality.

Conclusion Investigating longitudinal trajectory patterns of serial sex steroid and gonadotrophin measurements, the present study found no consistent associations with incident clinical CVD and all-cause mortality risk in elderly men from the community.

(Received 17 February 2012; returned for revision 25 April 2012; finally revised 06 June 2012; accepted 9 August 2012)

Correspondence: Robin Haring, Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Ferdinand-Sauerbruch Str., Greifswald 17475, Germany. Tel.: +49 3834 8619656;
Fax: +49 3834 865501;
E-mail: robin.haring@uni-greifswald.de

Introduction

Emerging data suggest that reduced circulating total testosterone (TT) concentrations in men are associated with increased cardiovascular risk and mortality.[1] Longitudinal epidemiological studies have consistently reported positive associations between low serum TT concentrations and hypertension,[2] metabolic syndrome,[3, 4] type 2 diabetes[5] and dyslipidaemia.[6] In contrast to these relations between low circulating TT concentrations in men and increased cardiometabolic risk factor burden, studies relating TT concentrations to incident cardiovascular disease (CVD) and mortality have remained inconclusive. While some longitudinal studies[7–10] reported an association between low TT concentrations and increased mortality risk, others like the Massachusetts Male Aging Study[11] or Caerphilly study[12] observed no such associations. Similarly, some longitudinal studies showed inverse associations between testosterone and cardiovascular events in elderly men,[13, 14] whereas a recent systematic review concluded that there was no association between testosterone and CVD risk in middle-aged men.[15] Furthermore, because a high CVD risk factor burden itself is associated with decreased TT concentrations,[16] there is a possibility of reverse causality when one considers the relations of circulating TT concentrations to CVD risk.

Analysing biomarker trajectories have previously been shown to offer additional insights into the assessment of the risk of CVD and mortality. For instance, in the Cardiovascular Health Study, a steep decline and extreme variability in dehydroepiandrosterone sulphate (DHEAS) concentrations were associated with increased mortality risk, whereas the baseline DHEAS concentration itself was not.[17] These findings raise the possibility that information regarding absolute hormone concentrations at a single point in time may be complemented by analysing serial measurements including hormone trajectories. Thus, we analysed up to four serial hormone measurements obtained over a 10-year period in elderly men from the original cohort of the Framingham Heart Study (FHS), to investigate longitudinal trajectory patterns of sex steroids and gonadotrophins. We hypothesized that low baseline and mean concentrations, as well as a steep decline and extreme variability in sex steroids and gonadotrophins are associated positively with incident clinical CVD and mortality.

Methods

Study population

The FHS is a longitudinal epidemiological study initiated in 1948 in Framingham, Massachusetts, to investigate risk factors for heart disease in a community-based epidemiological cohort design. The selection criteria and the study design of the FHS have been described previously.[18] Written informed consent was obtained at each examination and the Institutional Review Board of the Boston University Medical Center approved the study protocol. The present study focused on men who attended the biennial original cohort examination cycles 17 (1981–1983) through 20 (1987–1989) with two or more hormone measurements. Of the 1401 baseline attendants at examination 20, we excluded 866 women (our analyses were confined to men), 229 men with prevalent CVD, 32 men with missing hormone or covariate data and 20 men with less than two hormone measurements, yielding a final study sample of 254 men for the present investigation.

Measures

Single nonfasting blood samples were taken all-day from attendees at the original cohort's 17th–20th examination cycles between May 1981 and December 1989 and were stored at −20 °C until assays of sex hormones and gonadotrophins between October 1984 and March 1990. The maximum storage length (for only a portion of the samples) was about 5 years. Concentrations of serum TT, DHEAS, follicle-stimulating hormone (FSH), luteinizing hormone (LH) and total estradiol (EST) were measured using radioimmunoassays (Diagnostic Products Corp., Los Angeles, CA, USA) in the Boston University laboratory of Dr. Sawin. The interassay coefficients of variation in the Sawin laboratory were as follows: for TT, 11%; DHEAS, 11%; FSH, 5%; LH, 6%; and EST, 4%, as previously described.[19] The vast majority of the studied 254 men contributed four repeated measurements for TT (two measurements, 4; three, 41; four, 209), DHEAS (two measurements, 55; four, 199), FSH (two measurements, 4; three, 44; four, 206), LH (two measurements, 4; three, 42; four, 208) and EST (two measurements, 5; three, 44; four, 205).

Socio-demographic, behavioural and anthropometric characteristics as well as medical history, laboratory measures and medication use were assessed by standardized methods at the Heart Study examinations. Clinical CVD was assessed according to previously reported standardized criteria [including coronary heart disease (recognized or unrecognized myocardial infarction, angina pectoris, coronary insufficiency or coronary heart disease death), cerebrovascular disease (stroke or transient ischaemic attack) or congestive heart failure] and confirmed, together with information about vital status, by a three-physician adjudication panel with the aid of medical histories, physical examinations at the study clinic, hospitalization records and communication with personal physicians.[20]

Statistical analyses

Clinical CVD and mortality outcomes were assessed prospectively over 5- and 10-year follow-up periods starting from the 20th examination cycle (Fig. 1). To understand the nature of sampling bias because of nonattendance at serial exams or because of nonrandom measurement of hormones at serial exams, we compared clinical characteristics of men with two or more serial hormone measurements to those with less than two measurements using χ^2 tests for categorical data or two-sample *t*-tests for continuous data.

We used age- and multivariable-adjusted Cox proportional hazard regression models to relate baseline and mean hormone concentrations and their trajectories over time to clinical CVD incidence and all-cause mortality at the 5- and 10-year follow-up, by assessing (1) *baseline level:* sex steroid and gonadotrophin concentration from the first available baseline examination of

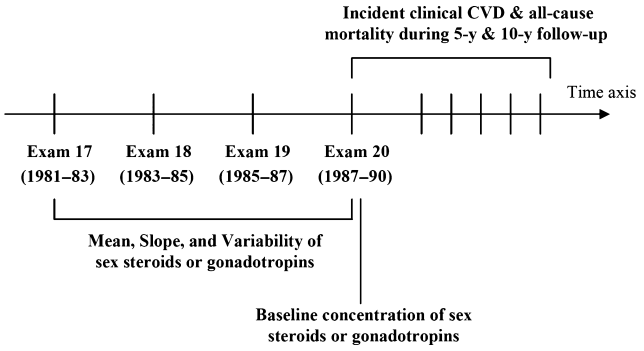


Fig. 1 Study design. Baseline concentrations of sex steroids and gonadotrophins, as well as covariates were measured at the 20th examination cycle of the Framingham Heart Study. From the 17th to 20th examination cycle, serial hormone measurements were used to assess their mean, slope and variability (separate analyses for each of the hormones). Incident clinical cardiovascular disease and all-cause mortality were analysed at the 5- and 10-year follow-up, respectively.

examination cycles 17–20 (most from the 20th examination cycle); (2) *mean:* average sex steroid and gonadotrophin concentration calculated in each individual using measures from at least two and up to four examination cycles; (3) *slope*: calculated using at least two available measures and the following regression formula: $y_i = a_i + \beta_i x_i$ where y_i is the predicted hormone concentration in the individual, a_i is the hormone concentration at the first examination, x_i is the number of years after the first examination and β_i is the hormone concentration's slope; and (4) *variability*: calculated using SAS RMSE (Root Mean Square Error; SAS Institute Inc., Cary, NC, USA) function to estimate the square root of the variance of the residuals based on at least three available measures ($n = 250$ for TT, FSH and LH; $n = 199$ for DHEAS; $n = 249$ for EST). We categorized the different predictor variables 1–4 into quartiles and included them as a continuous variable into the models. Effects are presented as hazard ratios (HR) and their 95% confidence intervals (95% CI).

Multivariable modelling included *covariates* measured in the 20th examination cycle: age, body mass index, smoking, total cholesterol, high-density lipoprotein cholesterol, type 2 diabetes, systolic blood pressure and antihypertensive medication. We confirmed that the proportional hazards assumption was satisfied by visual inspection of log-log plots and a test based on Schoenfeld residuals. As the probability of experiencing the event of interest (incident clinical CVD) may be altered by an alternative event (death), we conducted competing risk analyses for death in the Cox models relating sex steroids, gonadotrophins and their trajectories to incident clinical CVD.

To identify additional clinical and laboratory parameters predictive of incident CVD and all-cause mortality, we modelled baseline clinical and laboratory parameters (age, BMI, smoking status, systolic blood pressure, antihypertensive medication, type 2 diabetes, total cholesterol/high-density lipoprotein cholesterol (HDL-C) ratio, and sex steroid and gonadotrophin concentrations) with incident CVD and all-cause mortality at both 5- and 10-year follow-up in multivariable models. All statistical analyses were performed using SAS statistical software (SAS Institute Inc.). We used a minimal $P < 0{\cdot}005$ to indicate statistical significance because we evaluated five different hormones and two outcomes (0·05/10). The present study of 254 elderly men from the original cohort of the FHS had 80% power to detect a hazard ratio of 2·43 per quartile increment of any sex hormone/gonadotrophin for 10-year risk of incident CVD. For 10-year all-cause mortality, this study had 80% power to detect a hazard ratio of 1·76 per quartile increment.

Results

The baseline characteristics of our sample of older men are displayed in Table 1. Comparing participants included in the present analysis to the participants not included, we observed statistically significant differences with higher lipid levels and blood pressure, as well as a lower prevalence of diabetes and antihypertensive medication use in the first compared with the latter group. At the 5-year follow-up, we observed 56 incident clinical CVD cases and 42 deaths. We found no association between concentrations of sex steroids, gonadotrophins and their trajectories with incident clinical CVD (Table 2). Higher baseline TT concentrations were borderline significantly associated with lower all-cause mortality risk (HR per quartile increment, 0·74; 95% CI, 0·56–0·98; *P*-value, 0·034) in multivariable-adjusted, but not in age-adjusted models (HR per quartile increment, 0·81; 95% CI, 0·61–1·07; *P*-value, 0·134; Table 2). None of these results met the minimal $P < 0{\cdot}005$ criterion for statistical significance accounting

Table 1. Baseline characteristics of the 254 men in the study sample

Characteristic	
Age, years	75·5 ± 5·4
Body mass index, kg/m²	26·8 ± 3·8
Current smoker,%	7·1
Systolic blood pressure, mmHg	145·4 ± 19·7
Diastolic blood pressure, mmHg	78·2 ± 11·0
Antihypertensive medication,%	31·9
Type 2 diabetes,%	11·4
HDL cholesterol, mM	1·10 ± 0·30
Total cholesterol, mM	5·31 ± 0·95
Ratio total: HDL-C	5·2 ± 1·7

Sex steroids and gonadotrophins	TT, nmol/l	DHEAS, μmol/l	FSH, IU/l	LH, IU/l	EST, pmol/l
1. Baseline concentration	14·9 (11·4, 18·0)	2·14 (1·33, 3·18)	7·2 (4·9, 10·4)	8·8 (6·7, 12·3)	80 (51, 120)
2. Mean	14·9 (12·5, 17·7)	2·69 (1·85, 3·82)	8·4 (6·0, 11·9)	8·6 (7·0, 11·0)	86 (68, 110)
3. Slope	−1·0 (−2·4, 0·0)	−0·26 (−0·61, 0·06)	−0·5 (−1·2, 0·1)	0·2 (−0·5, 1·2)	−6 (−19, 6)
4. Variability	3·5 (2·1, 4·9)	0·49 (0·27, 0·83)	1·6 (0·8, 2·9)	2·1 (1·2, 2·9)	35 (22, 48)

HDL, high-density lipoprotein; TT, total testosterone; DHEAS, dehydroepiandrosterone sulphate; FSH, follicle-stimulating hormone; LH, luteinizing hormone; EST, total estradiol.
Data are percentages, mean ± SD or median (Q1, Q3).

for multiple testing. At the 10-year follow-up, we observed 92 incident clinical CVD cases and 104 deaths. However, repeat analyses at the 10-year follow-up also showed no consistent statistically significant associations between sex steroids, gonadotrophins and their trajectories with incident clinical CVD or all-cause mortality risk (Table 3). Adjusting the analyses for competing risks of incident clinical CVD at the 5- and 10-year follow-up showed no impact on the observed estimates (data not shown). Similarly, subgroup analyses stratified by BMI status (normal weight: BMI < 25 kg/m², 32·3%; overweight: BMI 25: <30 kg/m², 48·0%; obese: BMI: ≥30 kg/m², 19·7%) showed no association of sex steroids and gonadotrophins with incident 10-year CVD and all-cause mortality risk in any of the BMI strata (data not shown). Other clinical and laboratory predictors for incident CVD were age and smoking, and for all-cause mortality age and type 2 diabetes (data not shown).

Discussion

Leveraging the unique data set and design of the community-based FHS, the present study is the first to investigate longitudinal trajectory patterns of serial sex steroid and gonadotrophin measurements and their associations with 5- and 10-year risk of incident clinical CVD and all-cause mortality. We observed no consistent association of sex steroids, gonadotrophins and their trajectories with incident clinical CVD or all-cause mortality risk in 254 elderly men in the community.

In contrast to previous longitudinal studies[13, 14] and a recent meta-analysis[15] suggesting inverse associations between testosterone and cardiovascular events in elderly men over 70 years of age (with an estimated summary relative risk of 0·84 (0·76–0·92) per one SD increase in TT), the present and previous other longitudinal studies were not able to detect any statistically significant associations between TT concentrations and incident CVD.[21–23] In a 10-year follow-up of 2084 middle-aged men from the FHS, the results also suggested no association between TT or DHEAS concentrations and incident CVD.[21] Similarly, we found no consistent associations between baseline sex steroids, gonadotrophins and their longitudinal trajectory patterns with 5- and 10-year all-cause mortality risk. A recent meta-analyses summarized the reported associations between low TT concentrations and increased mortality risk and concluded that the considerable between-study heterogeneity related to study and participant characteristics limit the ability to provide valid sum-

Table 2. Sex steroids and gonadotrophins in men (in quartiles) and their trajectories associated with incident clinical cardiovascular disease (CVD) and all-cause mortality during 5-year follow-up

	TT	DHEAS	FSH	LH	EST
Incident CVD [HR per quartile increment (95% CI)]					
Baseline					
age-adjusted	0·92 (0·73, 1·16)	0·98 (0·77, 1·24)	1·11 (0·87, 1·42)	0·95 (0·75, 1·22)	0·87 (0·69, 1·11)
multivariable-adjusted	0·98 (0·77, 1·25)	1·00 (0·78, 1·27)	1·01 (0·79, 1·31)	0·85 (0·66, 1·11)	0·86 (0·68, 1·10)
Mean					
age-adjusted	0·98 (0·78, 1·24)	0·98 (0·77, 1·24)	1·11 (0·87, 1·41)	1·04 (0·82, 1·33)	0·91 (0·72, 1·15)
multivariable-adjusted	1·04 (0·81, 1·33)	1·00 (0·78, 1·27)	1·03 (0·81, 1·31)	0·95 (0·74, 1·23)	0·90 (0·71, 1·14)
Slope					
age-adjusted	0·94 (0·74, 1·20)	0·99 (0·78, 1·25)	0·82 (0·64, 1·04)	0·98 (0·78, 1·24)	0·99 (0·78, 1·25)
multivariable-adjusted	0·87 (0·68, 1·12)	1·00 (0·78, 1·28)	0·80 (0·63, 1·02)	0·92 (0·73, 1·17)	1·04 (0·82, 1·32)
Variability					
age-adjusted	0·99 (0·78, 1·25)	1·20 (0·93, 1·54)	1·00 (0·79, 1·26)	0·93 (0·74, 1·18)	0·87 (0·69, 1·10)
multivariable-adjusted	1·03 (0·82, 1·30)	1·29 (0·99, 1·69)	0·96 (0·76, 1·22)	0·92 (0·73, 1·17)	0·87 (0·69, 1·10)
All-cause mortality [HR per quartile increment (95% CI)]					
Baseline					
age-adjusted	0·81 (0·61, 1·07)	1·14 (0·87, 1·49)	0·90 (0·68, 1·20)	1·05 (0·78, 1·41)	0·98 (0·75, 1·29)
multivariable-adjusted	0·74 (0·56, 0·98)	1·11 (0·83, 1·49)	0·91 (0·67, 1·23)	0·99 (0·72, 1·36)	0·95 (0·72, 1·27)
Mean					
age-adjusted	0·98 (0·74, 1·30)	1·16 (0·88, 1·53)	1·05 (0·79, 1·40)	1·33 (0·98, 1·79)	0·98 (0·75, 1·28)
multivariable-adjusted	0·88 (0·67, 1·16)	1·19 (0·88, 1·61)	1·04 (0·77, 1·39)	1·20 (0·88, 1·65)	0·94 (0·71, 1·25)
Slope					
age-adjusted	0·79 (0·59, 1·05)	1·19 (0·90, 1·56)	0·82 (0·62, 1·08)	0·85 (0·65, 1·11)	0·91 (0·70, 1·18)
multivariable-adjusted	0·80 (0·59, 1·07)	1·07 (0·80, 1·43)	0·80 (0·61, 1·06)	0·80 (0·61, 1·05)	0·83 (0·63, 1·09)
Variability					
age-adjusted	1·17 (0·89, 1·55)	1·10 (0·80, 1·52)	1·00 (0·76, 1·32)	0·91 (0·70, 1·19)	0·94 (0·72, 1·22)
multivariable-adjusted	1·15 (0·88, 1·52)	1·09 (0·77, 1·53)	1·04 (0·79, 1·36)	0·94 (0·72, 1·22)	0·97 (0·74, 1·27)

HR, hazard ratio; TT, total testosterone; DHEAS, dehydroepiandrosterone sulphate; FSH, follicle-stimulating hormone; LH, luteinizing hormone; EST, total estradiol.

Multivariable models were adjusted for age, body mass index, smoking, total cholesterol, high-density lipoprotein cholesterol, type 2 diabetes, systolic blood pressure and antihypertensive medication. None of these results met the minimal $P < 0·005$ criterion for statistical significance accounting for multiple testing.

Table 3. Sex steroids and gonadotrophins in men (in quartiles) and their trajectories associated with incident clinical cardiovascular disease (CVD) and all-cause mortality during 10-year follow-up

	TT	DHEAS	FSH	LH	EST
Incident cardiovascular disease [HR per quartile increment (95% CI)]					
Baseline					
age-adjusted	0·98 (0·82, 1·18)	0·94 (0·78, 1·13)	1·09 (0·90, 1·32)	1·02 (0·84, 1·23)	0·91 (0·76, 1·10)
multivariable-adjusted	1·04 (0·86, 1·26)	0·97 (0·80, 1·18)	1·04 (0·86, 1·26)	0·95 (0·78, 1·16)	0·90 (0·75, 1·09)
Mean					
age-adjusted	0·99 (0·83, 1·20)	0·92 (0·76, 1·11)	1·07 (0·89, 1·29)	1·00 (0·83, 1·21)	0·91 (0·75, 1·09)
multivariable-adjusted	1·05 (0·87, 1·27)	0·95 (0·78, 1·16)	1·02 (0·85, 1·23)	0·94 (0·77, 1·14)	0·91 (0·75, 1·10)
Slope					
age-adjusted	0·98 (0·82, 1·18)	1·00 (0·84, 1·20)	0·93 (0·78, 1·12)	1·04 (0·87, 1·25)	0·99 (0·82, 1·19)
multivariable-adjusted	0·98 (0·81, 1·18)	0·99 (0·81, 1·19)	0·90 (0·75, 1·09)	1·00 (0·84, 1·21)	1·02 (0·85, 1·23)
Variability					
age-adjusted	0·90 (0·75, 1·09)	1·13 (0·93, 1·38)	0·97 (0·81, 1·16)	0·96 (0·80, 1·15)	0·96 (0·80, 1·15)
multivariable-adjusted	0·93 (0·78, 1·13)	1·18 (0·96, 1·45)	0·93 (0·77, 1·12)	0·96 (0·80, 1·15)	0·94 (0·78, 1·13)
All-cause mortality [HR per quartile increment (95% CI)]					
Baseline					
age-adjusted	0·93 (0·79, 1·11)	1·09 (0·92, 1·29)	0·96 (0·80, 1·14)	1·03 (0·86, 1·23)	1·04 (0·88, 1·24)
multivariable-adjusted	0·91 (0·76, 1·08)	1·12 (0·94, 1·34)	0·98 (0·81, 1·18)	1·02 (0·85, 1·23)	1·05 (0·88, 1·25)
Mean					
age-adjusted	1·01 (0·85, 1·20)	0·99 (0·83, 1·18)	0·98 (0·82, 1·18)	1·14 (0·96, 1·36)	1·07 (0·90, 1·27)
multivariable-adjusted	0·98 (0·82, 1·18)	1·03 (0·86, 1·25)	1·00 (0·83, 1·20)	1·10 (0·92, 1·33)	1·06 (0·89, 1·27)
Slope					
age-adjusted	0·88 (0·74, 1·06)	1·19 (1·00, 1·41)	1·05 (0·88, 1·25)	0·97 (0·82, 1·16)	1·01 (0·85, 1·20)
multivariable-adjusted	0·89 (0·74, 1·06)	1·13 (0·94, 1·35)	1·03 (0·86, 1·22)	0·97 (0·82, 1·15)	0·98 (0·81, 1·18)
Variability					
age-adjusted	1·03 (0·86, 1·22)	0·96 (0·79, 1·16)	1·01 (0·85, 1·20)	0·94 (0·79, 1·12)	1·04 (0·88, 1·23)
multivariable-adjusted	1·03 (0·87, 1·23)	1·00 (0·82, 1·22)	1·02 (0·86, 1·21)	0·95 (0·80, 1·13)	1·09 (0·92, 1·29)

TT, total testosterone; DHEAS, dehydroepiandrosterone sulphate; FSH, follicle-stimulating hormone; LH, luteinizing hormone; EST, total estradiol. Multivariable models were adjusted for age, body mass index, smoking, total cholesterol, high-density lipoprotein cholesterol, type 2 diabetes, systolic blood pressure and antihypertensive medication. None of these results met the minimal $P < 0{\cdot}005$ criterion for statistical significance accounting for multiple testing.

mary estimates or to draw any firm conclusions.[1] However, because of the advanced age and small size of the present study sample, competing risks of CVD and mortality risk and the limited statistical power may potentially explain the lack of association in our sample.

The present study is the first to analyse longitudinal trajectory patterns and their association with incident clinical CVD and mortality, but we did not observe any associations of mean hormone concentrations, their slope and variability with incident adverse events. In contrast, a previous study reported that a steep decline and extreme variability in DHEAS concentrations, but not the actual baseline level, were associated with increased all-cause mortality risk.[17] However, given the considerable potential for residual confounding (no adjustment was performed for BMI, BP or smoking) in that previous study and the negative finding for trajectory patterns as predictor of incident CVD and mortality in the present study, the postulated advantage of serial measurements over a single baseline measure is questionable.

The strengths of the present study include the use of four serial measurements of sex steroids and gonadotrophins from a well-characterized community-based sample of elderly men, enabling not only the analyses of baseline concentrations but also longitudinal trajectory patterns. For the proper interpretation of the present study, some important limitations must also be mentioned. Single nonfasting serum samples that were drawn through out the day at each individual examination were used to measure concentrations of sex steroids and gonadotrophins. Although we sought to limit artifactual changes in sex hormone concentrations across serial examinations by performing measurements in one central laboratory, following the same collection protocol, and using the same laboratory assays, measurement bias remains a possibility. However, a recent cross-sectional study of 95 elderly men aged 70 years also found no associations between sex steroid concentrations measured by liquid chromatography/tandem mass spectrometry and prevalent CVD.[24] Additionally, we did not measure sex hormone-binding globulin or androgen receptor polymorphisms (that may affect sex hormone concentrations and effects) or assess symptoms of low TT concentrations for the clinical diagnosis of primary hypogonadism or secondary hypogonadotrophic hypogonadism. Finally, the external validity or generalizability of our findings to other age groups or ethnicities is limited given our study sample of elderly Caucasian men.

In conclusion, the present study showed no association of baseline sex steroids, gonadotrophins and their longitudinal trajectory patterns with incident clinical CVD and all-cause

mortality in elderly men. Given the relatively small sample size, the advanced age of the study sample and the high rates of CVD and mortality, limited statistical power and competing risks hampered the detection of possible associations. However, the potential role of sex steroid concentrations in men as useful biomarkers of general health, cardiovascular risk factor burden and subclinical CVD progression warrants further investigation.

Disclosure statement

The present study was supported in part by contract NO1-HC25195 from the NHLBI.

References

1 Araujo, A.B., Dixon, J.M., Suarez, E.A. *et al.* (2011) Endogenous testosterone and mortality in men: a systematic review and meta-analysis. *The Journal of Clinical Endocrinology and Metabolism*, **96**, 3007–3019.
2 Torkler, S., Wallaschofski, H., Baumeister, S.E. *et al.* (2011) Inverse association between total testosterone concentrations, incident hypertension and blood pressure. *Aging Male*, **14**, 176–182.
3 Haring, R., Volzke, H., Felix, S.B. *et al.* (2009) Prediction of metabolic syndrome by low serum testosterone levels in men: results from the study of health in Pomerania. *Diabetes*, **58**, 2027–2031.
4 Kupelian, V., Page, S.T., Araujo, A.B. *et al.* (2006) Low sex hormone-binding globulin, total testosterone, and symptomatic androgen deficiency are associated with development of the metabolic syndrome in nonobese men. *The Journal of Clinical Endocrinology and Metabolism*, **91**, 843–850.
5 Vikan, T., Schirmer, H., Njolstad, I. *et al.* (2010) Low testosterone and sex hormone-binding globulin levels and high estradiol levels are independent predictors of type 2 diabetes in men. *European Journal of Endocrinology/European Federation of Endocrine Societies*, **162**, 747–754.
6 Haring, R., Baumeister, S.E., Völzke, H. *et al.* (2011) Prospective Association of Low Total Testosterone Concentrations with an Adverse Lipid Profile and Increased Incident Dyslipidemia. *European Journal of Cardiovascular Prevention and Rehabilitation*, **18**, 86–96.
7 Haring, R., Volzke, H., Steveling, A. *et al.* (2010) Low serum testosterone levels are associated with increased risk of mortality in a population-based cohort of men aged 20–79. *European Heart Journal*, **31**, 1494–1501.
8 Khaw, K.T., Dowsett, M., Folkerd, E. *et al.* (2007) Endogenous testosterone and mortality due to all causes, cardiovascular disease, and cancer in men: European prospective investigation into cancer in Norfolk (EPIC-Norfolk) Prospective Population Study. *Circulation*, **116**, 2694–2701.
9 Laughlin, G.A., Barrett-Connor, E. & Bergstrom, J. (2008) Low serum testosterone and mortality in older men. *The Journal of Clinical Endocrinology and Metabolism*, **93**, 68–75.
10 Tivesten, A., Vandenput, L., Labrie, F. *et al.* (2009) Low serum testosterone and estradiol predict mortality in elderly men. *The Journal of Clinical Endocrinology and Metabolism*, **94**, 2482–2488.
11 Araujo, A.B., Kupelian, V., Page, S.T. *et al.* (2007) Sex steroids and all-cause and cause-specific mortality in men. *Archives of Internal Medicine*, **167**, 1252–1260.
12 Smith, G.D., Ben-Shlomo, Y., Beswick, A. *et al.* (2005) Cortisol, testosterone, and coronary heart disease: prospective evidence from the Caerphilly study. *Circulation*, **112**, 332–340.
13 Ohlsson, C., Barrett-Connor, E., Bhasin, S. *et al.* (2011) High serum testosterone is associated with reduced risk of cardiovascular events in elderly men. The MrOS (Osteoporotic Fractures in Men) study in Sweden. *Journal of the American College of Cardiology*, **58**, 1674–1681.
14 Yeap, B.B., Hyde, Z., Almeida, O.P. *et al.* (2009) Lower testosterone levels predict incident stroke and transient ischemic attack in older men. *The Journal of Clinical Endocrinology and Metabolism*, **94**, 2353–2359.
15 Ruige, J.B., Mahmoud, A.M., De Bacquer, D. *et al.* (2010) Endogenous testosterone and cardiovascular disease in healthy men: a meta-analysis. *Heart (British Cardiac Society)*, **97**, 870–875.
16 Haring, R., Ittermann, T., Volzke, H. *et al.* (2010) Prevalence, incidence and risk factors of testosterone deficiency in a population-based cohort of men: results from the study of health in Pomerania. *Aging Male*, **13**, 247–257.
17 Cappola, A.R., O'Meara, E.S., Guo, W. *et al.* (2009) Trajectories of dehydroepiandrosterone sulfate predict mortality in older adults: the cardiovascular health study. *The Journals of Gerontology*, **64**, 1268–1274.
18 Dawber, T.R., Meadors, G.F. & Moore, F.E. Jr (1951) Epidemiological approaches to heart disease: the Framingham Study. *American Journal of Public Health and the Nation's Health*, **41**, 279–281.
19 Amin, S., Zhang, Y., Sawin, C.T. *et al.* (2000) Association of hypogonadism and estradiol levels with bone mineral density in elderly men from the Framingham study. *Annals of Internal Medicine*, **133**, 951–963.
20 Cupples, L.A., D'Agostino, R.B. & Kiely, D. (1987) The Framingham Study. An Epidemiological Investigation of Cardiovascular Disease, Section 34. Some Risk Factors Related to the Annual Incidence of Cardiovascular Disease and Death Using Pooled Repeated Biennial Measurements: Framingham Heart Study, 30 Year Follow-up. National Heart Lung, and Blood Institute, Bethesda, MD.
21 Arnlov, J., Pencina, M.J., Amin, S. *et al.* (2006) Endogenous sex hormones and cardiovascular disease incidence in men. *Annals of Internal Medicine*, **145**, 176–184.
22 Muller, M., van den Beld, A.W., Bots, M.L. *et al.* (2004) Endogenous sex hormones and progression of carotid atherosclerosis in elderly men. *Circulation*, **109**, 2074–2079.
23 Yarnell, J.W., Beswick, A.D., Sweetnam, P.M. *et al.* (1993) Endogenous sex hormones and ischemic heart disease in men. The Caerphilly prospective study. *Arteriosclerosis and Thrombosis*, **13**, 517–520.
24 Naessen, T., Sjogren, U., Bergquist, J. *et al.* (2010) Endogenous steroids measured by high-specificity liquid chromatography-tandem mass spectrometry and prevalent cardiovascular disease in 70-year-old men and women. *The Journal of Clinical Endocrinology and Metabolism*, **95**, 1889–1897.

International Journal of Andrology

international journal of andrology ISSN 0105-6263

ORIGINAL ARTICLE

Clinical correlates of sex steroids and gonadotropins in men over the late adulthood: the Framingham Heart Study

R. Haring,*§ V. Xanthakis,*†¶ A. Coviello,*‡ L. Sullivan,† S. Bhasin,‡ H. Wallaschofski,§ J. M. Murabito*¶ and R. S. Vasan*¶

*Preventive Medicine & Epidemiology Section and Section of General Internal Medicine, †Department of Biostatistics, ‡Section of Endocrinology, Diabetes, and Nutrition, Boston University School of Medicine, Boston, MA, USA, §Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany, and ¶National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA

Keywords:
ageing male, Framingham Heart Study, gonadotropins, sex steroids, testosterone

Correspondence:
Robin Haring, PhD,
Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Ferdinand-Sauerbruch-Strasse, Greifswald, D-17475; Germany.
E-mail: robin.haring@uni-greifswald.de and Ramachandran S. Vasan, MD, DM, FACC, FAHA, Preventive Medicine & Epidemiology Section and Section of General Internal Medicine, Boston University School of Medicine, 761 Harrison Avenue, Harrison Court, Boston, Massachusetts 02118; USA.
E-mail: vasan@bu.edu

Received 7 November 2011; revised 2 April 2012; accepted 10 April 2012

doi:10.1111/j.1365-2605.2012.01285.x

Summary

Low serum concentrations of sex steroids and gonadotropins in men have been associated with increased cardiometabolic risk and mortality, but the clinical correlates of these hormones in men over late adulthood are less clearly understood. We analysed up to five serial measurements of total testosterone (TT), dehydroepiandrosterone sulphate (DHEAS), follicle stimulating hormone (FSH), luteinizing hormone (LH) and total estradiol (EST) in older men in the original cohort of the Framingham Heart Study to determine the short- (2-years; 1,165 person-observations in 528 individuals) and long-term (up to 10-years follow-up; 2520 person-observations in 835 individuals with mean baseline age: 71.2 years) clinical correlates of these sex steroids and gonadotropins using multilevel modelling and Generalized Estimating Equations. Age, body mass index and pre-existing type 2 diabetes were inversely related to long-term TT concentrations, whereas higher systolic blood pressure showed a positive association. Furthermore, age and pre-existing cardiovascular disease (CVD) were inversely associated and HDL cholesterol concentrations positively associated with long-term DHEAS concentrations respectively. Analyses of short-term changes revealed age was inversely related to DHEAS, but positively related to FSH and LH concentrations. Our community-based study identified modifiable correlates of decreasing TT and DHEAS concentrations in elderly men, suggesting that maintenance of a low CVD risk factor burden may mitigate the age-related decline of these hormones over the late adulthood.

Introduction

Low serum concentrations of total testosterone (TT) in men have been associated with increased cardiometabolic risk factor burden including a greater prevalence of dyslipidemia (Haring *et al.*, 2011), hypertension (Torkler *et al.*, 2011), metabolic syndrome (MetS) (Kupelian *et al.*, 2006; Haring *et al.*, 2009), type 2 diabetes (Vikan *et al.*, 2010; Schipf *et al.*, 2011) and atherosclerosis (Svartberg *et al.*, 2006; Vikan *et al.*, 2009), as well as mortality risk (Araujo *et al.*, 2007; Laughlin *et al.*, 2008; Haring *et al.*, 2010b). In addition, prostate cancer patients who undergo long-term androgen deprivation therapy are at greater risk of developing dyslipidemia, insulin resistance, hyperglycaemia and MetS, suggesting potentially beneficial effects of endogenous TT on cardiovascular disease (CVD) risk factor burden (Hakimian *et al.*, 2008). But prospective studies relating sex steroids to incident CVD have yielded inconsistent results (Muller *et al.*, 2003b) and an initial trial of testosterone therapy was discontinued after it showed adverse cardiovascular events in the treatment group (Basaria *et al.*, 2010). Furthermore, the interpretation of temporality of associations is challenging based on cross-sectional data because of the possibility of reverse causality. Thus, low TT concentrations are associated with incident MetS (Haring *et al.*, 2009) or type 2 diabetes (Vikan *et al.*, 2010; Schipf *et al.*, 2011), whereas these conditions are also associated with decreased TT concentrations (Laaksonen *et al.*, 2005; Corona *et al.*, 2011). Although it has been widely observed that TT and

other sex steroids in men decline with age (Gray *et al.*, 1991; Harman *et al.*, 2001; Feldman *et al.*, 2002; Svartberg *et al.*, 2003), other contributing factors have also been proposed (Andersson *et al.*, 2007; Travison *et al.*, 2007b; Haring *et al.*, 2010a), suggesting a likely multifactorial basis for the decline in circulating TT concentrations in ageing men (Snyder, 2008).

Overall, there is a gap in our current knowledge about whether low TT serum concentrations in men are causal or the consequence of CVD. Also, the clinical correlates of longitudinal tracking of TT concentrations in men from the general population have been less well studied (Snyder, 2008). Accordingly, we examined the longitudinal profile of sex steroids and gonadotropins in older men from the Framingham Heart Study (FHS) using up to five serial hormone measurements over a 10-year follow-up period, and evaluated the clinical correlates. Specifically, we assessed both the short-term (2-year) change and the long-term (10-year) tracking for these steroid and non-steroid hormones.

Methods

Study population

The FHS is a longitudinal epidemiological study that was initiated in 1948 in Framingham, Massachusetts, to investigate risk factors for heart disease in the community. The selection criteria and study design of the original FHS cohort have been described previously (Dawber *et al.*, 1951). Written informed consent was obtained from attendees at each examination and the Institutional Review Board of the Boston University Medical Center approved the study protocol. The present study focused on men who attended at least two of the biennial examination cycles 17 (1981–1983) through 21 (1990–1992). We excluded observations with missing outcome or covariate information. We did not use information regarding testosterone therapy or hypogonadism as additional exclusion criteria as such data were not systematically collected at these examinations. Correlates of short-term change in sex steroids and gonadotropins (over any 2-year period) were investigated pooling 1165 person-observations from 528 unique participants who attended two consecutive examinations. Long-term tracking of sex steroids and gonadotropins (over the course of 10 years) was analysed in 835 unique participants (2520 person-observations) (Fig. 1).

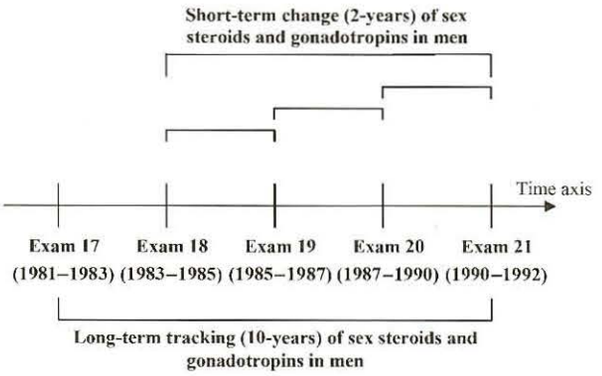


Figure 1 Study design. Short-term change in sex steroids and gonadotropins was evaluated in 528 men (1165 person-observations). Long-term tracking of sex steroids and gonadotropins was performed in 835 men who attended up to five serial examination cycles (2520 person-observations).

Laboratory measurements

Single non-fasting blood samples were obtained throughout the day from attendees at the original cohort's 17th – 21st examination cycles between May 1981 and December 1989, and the biosamples were stored at −20 °C until measurement of sex hormones and gonadotropins between October 1984 to March 1990. The maximum storage length (for only a portion of the samples) was about 5 years. Concentrations of serum TT, DHEAS, follicle stimulating hormone (FSH), luteinizing hormone (LH) and total estradiol (EST) were measured using radioimmunoassays (Diagnostic Products Corp., Los Angeles, CA, USA) in the Boston University laboratory of Dr. Sawin. The interassay coefficients of variation in the Sawin laboratory were: for TT, 11%; DHEAS, 11%; FSH, 5%; LH, 6%; and EST, 4%, as previously described (Amin *et al.*, 2000).

Clinical correlates

At each FHS examination, socio-demographic and behavioural characteristics (such as age, smoking status etc.), as well as medical history and medication use were assessed using standardized interviews. Height and weight were measured and body mass index (BMI) was calculated (kg/m^2). Blood pressure (BP) was measured twice in the left arm of the seated subject with a mercury column sphygmomanometer. The average of the two readings was used as the exam BP, and hypertension was defined as systolic BP ≥140 mmHg or a diastolic BP ≥90 mmHg or the self-reported use of antihypertensive medications. Type 2 diabetes was defined by a non-fasting glucose >200 mg/dL, or self-reported use of insulin or oral hypoglycemic medications. Plasma total cholesterol and high-density lipoprotein (HDL) cholesterol concentrations were measured using standard enzymatic methods, as previously described (McNamara & Schaefer, 1987). Pre-existing CVD was defined according to previously reported standardized FHS criteria (including coronary heart disease, cerebrovascular disease, intermittent

claudication or congestive heart failure) and confirmed with the aid of medical histories, physical examinations at the study clinic, hospitalization records and communication with personal physicians as previously described (Kannel *et al.*, 1979).

Statistical analyses

Correlates of short-term change of sex steroids and gonadotropins in men

We naturally logarithmically transformed serum DHEAS, FSH, LH and EST concentrations to normalize their distributions. Generalized Estimating Equations (GEE) were used to determine clinical correlates of short-term change in sex steroids and gonadotropins during a 2-year follow-up period (Fig. 1). In these multivariable models, we related hormone concentrations (dependent variable; each hormone considered individually) to the following clinical covariates (independent variables): age, BMI, systolic BP, antihypertensive treatment, smoking status, type 2 diabetes, total cholesterol, HDL cholesterol and pre-existing CVD. These variables were chosen based on their previously reported associations with steroid and non-steroid hormone concentrations in the literature (Travison *et al.*, 2007a; Haring *et al.*, 2010a). To account for potential temporal trends in sex steroids and gonadotropins across different examination cycles, we carefully adjusted all statistical analyses for examination cycle. Furthermore, we provided boxplots for the median sex steroid and gonadotropin concentrations over time for a reference age group of individuals aged 65–75 years at each examination cycle, indicating no large variability across the different examination cycles (Supplemental Fig. S1). Interaction terms between age and each clinical covariate were investigated using multivariable models. To analyse patterns of missingness we compared characteristics of the study sample according to the number of missing examinations (Supplemental Tables S1 and S2).

Effects are presented as regression coefficients and their corresponding 95% confidence interval (95% CI).

Correlates of long-term tracking of sex steroids and gonadotropins in men

We performed multilevel statistical modelling (SAS PROC MIXED; using an unstructured correlation matrix) to identify clinical correlates of long-term tracking of sex hormone concentrations over a 10-year period. Accommodating participants with missing data at some of the serial examinations over the 10-year follow-up period, this analytical approach allows the maximization of the available number of observations in a longitudinal study design and accounts for a hierarchical data structure that varies on the individual level. Multivariable models incorporated the same set of covariates and interaction terms as used in the short-term analyses described above. While change in covariates is captured in the long-term follow-up analysis, the short-term analysis used "baseline" covariate data and only the hormone data from the baseline and the subsequent examination cycle. For the paired 2-year data, participants had to attend consecutive examinations while for the 10-year data no such sequential follow-up was mandated (not every individual has to attend every examination), which is why the latter approach provided more person observations. We examined and detected statistically significant random age effects for DHEAS, FSH, LH and EST, and modelled age as a random and a fixed effect accordingly. We also examined non-linear age effects, by including the 'squared age' term into the regression models and assessing the *p*-value for the term. As the 'squared age' term was not statistically significant in any of the models, we did not include it in further regression modelling.

Graphical representation of long-term sex steroids in men

To illustrate the impact of an individual's risk factor burden on long-term TT concentrations, graphical displays show the adjusted (for all significant variables from the long-term analyses) mean TT concentrations with increasing age stratified by 'high' vs. 'low' risk factor burden. The covariates used to define high and low risk factor status (age, BMI, systolic BP, type 2 diabetes and smoking) were selected on the basis of their statistically significant associations with TT concentrations in the final regression model. All statistical analyses were performed using SAS statistical software (SAS Institute Inc., Cary, NC, USA), and the figures were generated using Excel (Microsoft Office 2003, Redmond, WA, USA).

Results

The baseline characteristics of our sample of older men are displayed in Table 1.

Correlates of short-term change in sex steroids and gonadotropins in men

We identified age as the main correlate of short-term change in sex steroids and gonadotropins (Table 2), inversely associated with change in log-DHEAS (β per 1 year increase in age, −0.007; 95% CI, −0.012 to −0.002) and positively associated with change in log-FSH and log-LH concentrations (β per 1 year increase in age, 0.005; 95% CI, 0.001–0.008 and 0.007; 95% CI, 0.002–0.012, respectively). BMI was inversely associated with change in TT concentrations (β per one unit increase in

Table 1 Baseline characteristics of the study population with 10-year follow-up[a]

Characteristic	
Age, years	71.2 ± 6.4
Body mass index, kg/m^2	26.6 ± 3.8
Current smoker, %	16.5
Systolic blood pressure, mmHg	141.6 ± 18.5
Diastolic blood pressure, mmHg	78.0 ± 10.0
Antihypertensive medication, %	41.2
Hypertension, %	68.4
Type 2 diabetes, %	12.7
HDL cholesterol, mg/dL	44.3 ± 13.6
Total cholesterol, mg/dL	216.6 ± 37.4
Ratio total : HDL cholesterol, mg/dL	5.3 ± 1.6
Prevalence of cardiovascular disease, %	34.3
Total testosterone, ng/mL	4.9 (3.9; 5.9)
DHEAS, mg/dL	91.2 (56.8; 137.0)
FSH, IU/L	9.0 (6.2; 13.4)
Luteinizing hormone, IU/L	8.6 (6.2; 12.0)
Total estradiol, pg/mL	30.2 (21.3; 39.2)

Data are percentages, mean ± SD, or median (Q1; Q3).
[a]Baseline characteristics are presented for the sample with the largest available data for sex steroids and gonadotropins: total testosterone, N = 834; dehydroepiandrosterone sulphate (DHEAS), N = 657, follicle stimulating hormone (FSH), N = 835, luteinizing hormone, N = 835; estradiol, N = 834. Values for these hormones are reported based on availability of each hormone.

Table 2 Directionality of correlates of short-term change and long-term tracking of sex steroids and gonadotropins in men based on multivariable analyses

Correlates	Short-term change (over any 2-year period)	Long-term tracking (complete 10-year period)
Age	↓ DHEAS, ↑ FSH, ↑ LH	↓ TT, ↓ DHEAS, ↑ LH
Body mass index	↓ TT	↓ TT
Current smoking		↑ LH
Systolic blood pressure	↑ DHEAS	↑ TT, ↑ EST
Antihypertensive medication		↑ EST
Type 2 diabetes		↓ TT
HDL cholesterol		↑ DHEAS
Pre-existing cardiovascular disease		↓ DHEAS

TT, total testosterone; DHEAS, dehydroepiandrosterone sulphate; FSH, follicle-stimulating hormone; LH, luteinizing hormone; EST, total estradiol.

BMI, −0.02 ng/mL; 95% CI, −0.04 to −0.007 ng/mL) and systolic BP was positively associated with change in log-DHEAS concentrations (β per 10 mmHg increase in systolic BP, 0.02; 95% CI, 0.01–0.03).

Correlates of long-term tracking of sex steroids and gonadotropins in men

Age, BMI, smoking, systolic BP, hypertension treatment, type 2 diabetes, HDL cholesterol concentrations and pre-existing CVD were identified as significant correlates of changes in sex steroids and gonadotropins over the 10-year follow-up period (Tables 2 and 3). With regard to long-term TT concentrations, we found inverse associations with age, BMI and type 2 diabetes, and a positive association with systolic BP respectively. The interaction between age and smoking was statistically significant, indicating that the inverse effect of age on TT concentrations varies according to smoking status (smokers have a greater age-related decline in TT concentrations compared with non-smokers). Consistent with the analyses of short-term change, age was inversely associated with log-DHEAS concentrations and positively associated with log-LH concentrations in long-term analyses. Furthermore, HDL cholesterol concentrations were positively and pre-existing CVD inversely associated with log-DHEAS concentrations in long-term analyses. Other correlates and their directionality are listed in Tables 2 and 3. Finally, Fig. 2 illustrates the conjoint effect of multiple risk factors (categorized into two groups for example) on TT concentrations over the 10-year period.

Discussion

Our longitudinal study identified various clinical correlates of short-term changes and long-term tracking of sex steroids and gonadotropins in older men respectively. Analyses of short-term changes revealed age as inversely related to DHEAS, but positively related to FSH and LH concentrations. Age, BMI and type 2 diabetes were inversely related to long-term TT concentrations, whereas higher systolic BP showed a positive association. Thus, our study offers important insights into the correlates of long-term progression of sex steroids and gonadotropins in community-dwelling older men over their late adulthood.

Our study confirms previous longitudinal data reporting age as the main correlate of sex steroids and gonadotropins in men (Gray *et al.*, 1991; Morley *et al.*, 1997; Leifke *et al.*, 2000; Harman *et al.*, 2001; Muller *et al.*, 2003a; Lapauw *et al.*, 2008; Wu *et al.*, 2008; Cappola *et al.*, 2009), showing inverse associations with TT (Harman *et al.*, 2001) and DHEAS (Cappola *et al.*, 2009) concentrations, and positive associations with FSH (Morley *et al.*, 1997; Lapauw *et al.*, 2008) and LH (Morley *et al.*, 1997; Lapauw *et al.*, 2008; Wu *et al.*, 2008) concentrations. The fact that age was the only consistent correlate in our analyses of short-term changes may be explained by the

Table 3 Correlates of long-term tracking of sex steroids and gonadotropins in men based on multivariable analyses

	TT	DHEAS	FSH	LH	EST
Correlates	Beta coef. (95% CI)	Beta coef. (95% CI)	Beta coef. (95% CI)	Beta coef. (95% CI)	Beta coef. (95% CI)
Age	−0.04 (−0.05, −0.02) *	−0.03 (−0.04, −0.02) *	−0.01 (−0.04, 0.02)	0.03 (0.02, 0.04) *	
Body mass index	−0.04 (−0.06, −0.02) *				
Current smoking	0.17 (−0.06, 0.41)			0.09 (0.02, 0.17) *	
Systolic blood pressure	0.004 (0.001, 0.007) *		−0.0004 (−0.002, 0.0007)		0.002 (0.001, 0.004) *
Antihypertensive medication					0.08 (0.02, 0.14) *
Type 2 diabetes	−0.23 (−0.45, −0.02) *				
HDL cholesterol,		0.003 (0.0002, 0.006) *			
Pre-existing CVD		−0.10 (−0.17, −0.03) *			
Age × systolic Blood pressure			0.0003 (0.0001, 0.0005) *		
Age × smoking	−0.06 (−0.09, −0.03) *				

*$p < 0.05$.

CVD, cardiovascular disease, TT, total testosterone; DHEAS, dehydroepiandrosterone sulphate; FSH, follicle-stimulating hormone; LH, luteinizing hormone; EST, total estradiol.

Any non-significant covariates were retained in the model if they contributed to a significant interaction term.

Age was centred at the mean of all participants at all exams (73 years) to reduce multicollinearity between regression coefficients.

DHEAS, EST, FSH, and LH were naturally log transformed, therefore the coefficient (coef.) indicates an e^{β}-fold change in the respective sex steroid or gonadotropin. For example: β for age (log-DHEAS) = −0.03 → $e^{-0.03}$ = 0.97 fold decrease in DHEAS concentration per year increase. TT was used untransformed, therefore a one-unit increase in body mass index resulted in a 0.04 ng/mL decrease in TT concentration.

The effect of variables that are included in statistically significant interaction terms needs to be interpreted taking into account the respective interaction terms. For example, the effect of age on TT depends on smoking status (because of the presence of the statistically significant age × smoking interaction term in the model), therefore a 1-year increase in age results in a 0.1 ng/mL decrease in TT among smokers and a 0.04 ng/mL decrease in TT among non-smokers.

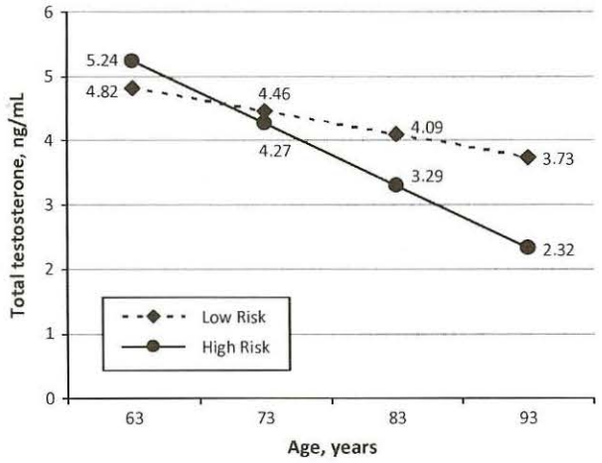


Figure 2 Adjusted mean total testosterone concentrations with increasing age. Long-term tracking of total testosterone concentrations in men with low and high CVD risk factor burden. Low risk factor burden: body mass index, 25 kg/m^2; systolic blood pressure, 130 mmHg; non-smoker; no pre-existing type 2 diabetes; high risk factor burden: body mass index, 30 kg/m^2; systolic blood pressure, 150 mmHg; smoker; pre-existing type 2 diabetes.

relatively short follow-up time of 2 years, limiting our ability to elucidate the effects of other covariates on sex steroids and gonadotropins in men. But besides ageing alone, the clinical correlates of long-term TT concentrations identified in this study are consistent with previous studies reporting visceral obesity (Derby *et al.*, 2006; Mohr *et al.*, 2006; Travison *et al.*, 2007a; Wu *et al.*, 2008), MetS (Laaksonen *et al.*, 2005; Haring *et al.*, 2010a), type 2 diabetes (Vermeulen *et al.*, 1996; Haring *et al.*, 2010a) and comorbidity (Svartberg *et al.*, 2003; Derby *et al.*, 2006; Haring *et al.*, 2010a) related to decreased TT concentrations. In contrast to previous studies reporting an inverse association between TT concentrations and BP among middle-aged men (Barrett-Connor & Khaw, 1988; Yarnell *et al.*, 1993; Torkler *et al.*, 2011), the present study is the first to report a positive association between systolic BP and TT among elderly men, whereas others did not observe any association at all (Zmuda *et al.*, 1997; Khaw *et al.*, 2007). However, these inconsistencies may relate to differences in study design, study sample, characteristics of the study population including age and underlying comorbidity or confounders adjusted for. Concerning DHEAS concentrations, a 9-year follow-up of 989 older men and women (mean age 85.2 years) from the Cardiovascular Health Study also showed that pre-existing CVD was associated with greater incident DHEAS decline (Sanders *et al.*, 2010). Furthermore, our finding of a positive association between HDL cholesterol and DHEAS concentrations in

long-term analyses was consistent with previous results from small cross-sectional studies (Haffner *et al.*, 1993; Yasui *et al.*, 2008).

Special attention belongs to our finding that BMI showed an inverse association with TT concentrations. It has been previously shown that a 4–5 kg/m^2 increase in BMI is associated with declines in TT concentrations comparable to that associated with approximately 10 years of ageing (Travison *et al.*, 2007a). Furthermore, we observed, consistent with previous longitudinal studies that pre-existing type 2 diabetes (Laaksonen *et al.*, 2005; Corona *et al.*, 2011) and pre-existing CVD (Sanders *et al.*, 2010) are associated with lower TT and DHEAS concentrations. Given the previously shown associations of low TT and DHEAS concentrations with incident type 2 diabetes (Schipf *et al.*, 2011; Vikan *et al.*, 2010) and pre-existing CVD (Barrett-Connor *et al.*, 1986) respectively, the present results provide further evidence for bidirectional influences between sex steroids and chronic diseases (Yeap, 2009).

Our reported positive associations between age and LH concentrations and also between smoking and LH concentrations confirms previous cross-sectional findings from the European Male Ageing Study among 3,220 men aged 40–79 years (Wu *et al.*, 2008) and extends them using serial longitudinal observations. As LH stimulates the Leydig cells to secrete testosterone and the quantity of testosterone secreted increases approximately in direct proportion to the amount of LH available, the effects of smoking are tightly linked for both steroid and non-steroid hormones in men (Mendelson *et al.*, 2003); which could possibly explain the observed positive associations between current smoking, LH (Wu *et al.*, 2008) and TT (Vermeulen *et al.*, 1996; Wu *et al.*, 2008).

Strengths and limitations

The strengths of the present investigation include the use of multilevel modelling in a unique community-based sample of older men with up to five serial measurements of sex steroids and gonadotropins over a 10-year period. Some important limitations also have to be mentioned. First, we used single serum samples and radioimmunoassays to measure circulating steroid and non-steroid hormone concentrations. However, because changes were related to baseline measures, the necessity of repeated testing at single time points to characterize borderline hormone concentrations does not apply to our study aims (Rosner *et al.*, 2007). Furthermore, we sought to limit artifactual changes in sex steroids and gonadotropins across serial examinations by performing measurements in one central laboratory, following the same collection protocol, and using the same laboratory assays for serum samples that were stored at −70 °C. In addition, we carefully adjusted all our analyses for "examination cycle" to account for potential measurement bias. Second, we did not measure sex-hormone binding globulin and were therefore not able to examine its correlates and interplay with the identified correlates of steroid and non-steroid hormone concentrations in men. Third, the external validity or generalizability of our findings to other populations, age groups or ethnicities is limited because of a community-based study sample of predominantly white older Caucasian men.

Conclusions

Given the accumulating evidence suggesting that low sex steroid concentrations in men may be associated with greater cardiometabolic risk, it is crucial to characterize their correlates. But the identified correlates influencing long-term steroid and non-steroid hormone concentrations in older men also constitute major cardiometabolic risk factors associated with CVD onset and progression. Thus, reverse causation might explain some of the observed associations in the literature, wherein adverse cardiometabolic risk factor profiles influence (lower) TT concentrations, which in turn may affect cardiometabolic risk factor burden. To further elucidate the potential role of low TT as a causal CVD risk factor (and the direction of causality), future research from large randomized controlled clinical trials of testosterone replacement therapy is needed. However, the present findings assert that prevention strategies should focus on health maintenance including a low cardiometabolic risk factor burden, instead of testosterone replacement therapy for improving CV health and lowering CVD risk.

Disclosures

None.

Sources of funding

This work was supported by the National Heart Blood and Lung Institute (contract N01-HC-25195).

References

Amin S, Zhang Y, Sawin CT, Evans SR, Hannan MT, Kiel DP, Wilson PW & Felson DT. (2000) Association of hypogonadism and estradiol levels with bone mineral density in elderly men from the Framingham study. *Ann Intern Med* 133, 951–963.

Andersson AM, Jensen TK, Juul A, Petersen JH, Jorgensen T & Skakkebaek NE. (2007) Secular decline in male testosterone and sex hormone binding globulin serum levels in Danish population surveys. *J Clin Endocrinol Metab* 92, 4696–4705.

Araujo AB, Kupelian V, Page ST, Handelsman DJ, Bremner WJ & McKinlay JB. (2007) Sex steroids and all-cause and cause-specific mortality in men. *Arch Int Med* 167, 1252–1260.

Barrett-Connor E & Khaw KT. (1988) Endogenous sex hormones and cardiovascular disease in men. A prospective population-based study. *Circulation* 78, 539–545.

Barrett-Connor E, Khaw KT & Yen SS. (1986) A prospective study of dehydroepiandrosterone sulfate, mortality, and cardiovascular disease. *New England J Med* 315, 1519–1524.

Basaria S, Coviello AD, Travison TG, Storer TW, Farwell WR, Jette AM *et al.* (2010) Adverse events associated with testosterone administration. *New England J Med* 363, 109–122.

Cappola AR, O'Meara ES, Guo W, Bartz TM, Fried LP & Newman AB. (2009) Trajectories of dehydroepiandrosterone sulfate predict mortality in older adults: the cardiovascular health study. *J Gerontol* 64, 1268–1274.

Corona G, Monami M, Rastrelli G, Aversa A, Sforza A, Lenzi A, Forti G, Mannucci E & Maggi M. (2011) Type 2 diabetes mellitus and testosterone: a meta-analysis study. *Int J Androl* 34, 528–540.

Dawber TR, Meadors GF & Moore FE, Jr. (1951) Epidemiological approaches to heart disease: the Framingham Study. *Am J Pub Health Nation's Health* 41, 279–281.

Derby CA, Zilber S, Brambilla D, Morales KH & McKinlay JB. (2006) Body mass index, waist circumference and waist to hip ratio and change in sex steroid hormones: the Massachusetts Male Ageing Study. *Clin Endocrinol (Oxf)* 65, 125–131.

Feldman HA, Longcope C, Derby CA, Johannes CB, Araujo AB, Coviello AD, Bremner WJ & McKinlay JB. (2002) Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts male aging study. *J Cli Endocrinol Metab* 87, 589–598.

Gray A, Feldman HA, McKinlay JB & Longcope C. (1991) Age, disease, and changing sex hormone levels in middle-aged men: results of the Massachusetts Male Aging Study. *J Clin Endocrinol Metab* 73, 1016–1025.

Haffner SM, Mykkanen L, Valdez RA & Katz MS. (1993) Relationship of sex hormones to lipids and lipoproteins in nondiabetic men. *J Clin Endocrinol Metab* 77, 1610–1615.

Hakimian P, Blute M, Jr, Kashanian J, Chan S, Silver D & Shabsigh R. (2008) Metabolic and cardiovascular effects of androgen deprivation therapy. *BJU Int* 102, 1509–1514.

Haring R, Volzke H, Felix SB, Schipf S, Dorr M, Rosskopf D, Nauck M, Schofl C & Wallaschofski H. (2009) Prediction of metabolic syndrome by low serum testosterone levels in men: results from the study of health in Pomerania. *Diabetes* 58, 2027–2031.

Haring R, Ittermann T, Volzke H, Krebs A, Zygmunt M, Felix SB, Grabe HJ, Nauck M & Wallaschofski H. (2010a) Prevalence, incidence and risk factors of testosterone deficiency in a population-based cohort of men: results from the study of health in Pomerania. *Aging Male* 13, 247–257.

Haring R, Volzke H, Steveling A, Krebs A, Felix SB, Schofl C, Dorr M, Nauck M & Wallaschofski H. (2010b) Low serum testosterone levels are associated with increased risk of mortality in a population-based cohort of men aged 20-79. *Eur Heart J* 31, 1494–1501.

Haring R, Baumeister SE, Völzke H, Dorr M, Felix SB, Kroemer HK, Nauck M & Wallaschofski H. (2011) Prospective association of low total testosterone concentrations with an adverse lipid profile and increased incident dyslipidemia. *Eur J Cardiovasc Prev Rehabil* 18, 86–96.

Harman SM, Metter EJ, Tobin JD, Pearson J & Blackman MR. (2001) Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. *J Clin Endocrinol Metab* 86, 724–731.

Kannel WB, Feinleib M, McNamara PM, Garrison RJ & Castelli WP. (1979) An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 110, 281–290.

Khaw KT, Dowsett M, Folkerd E, Bingham S, Wareham N, Luben R, Welch A & Day N. (2007) Endogenous testosterone and mortality due to all causes, cardiovascular disease, and cancer in men: European prospective investigation into cancer in Norfolk (EPIC-Norfolk) Prospective Population Study. *Circulation* 116, 2694–2701.

Kupelian V, Page ST, Araujo AB, Travison TG, Bremner WJ & McKinlay JB. (2006) Low sex hormone-binding globulin, total testosterone, and symptomatic androgen deficiency are associated with development of the metabolic syndrome in nonobese men. *J Clin Endocrinol Metab* 91, 843–850.

Laaksonen DE, Niskanen L, Punnonen K, Nyyssonen K, Tuomainen TP, Valkonen VP & Salonen JT. (2005) The metabolic syndrome and smoking in relation to hypogonadism in middle-aged men: a prospective cohort study. *J Clin Endocrinol Metab* 90, 712–719.

Lapauw B, Goemaere S, Zmierczak H, Van Pottelbergh I, Mahmoud A, Taes Y, De Bacquer D, Vansteelandt S & Kaufman JM. (2008) The decline of serum testosterone levels in community-dwelling men over 70 years of age: descriptive data and predictors of longitudinal changes. *Eu J Endocrinol / Eur Federation of Endocrine Soc* 159, 459–468.

Laughlin GA, Barrett-Connor E & Bergstrom J. (2008) Low serum testosterone and mortality in older men. *J Clin Endocrinol Metab* 93, 68–75.

Leifke E, Gorenoi V, Wichers C, Von Zur Muhlen A, Von Buren E & Brabant G. (2000) Age-related changes of serum sex hormones, insulin-like growth factor-1 and sex-hormone binding globulin levels in men: cross-sectional data from a healthy male cohort. *Clin Endocrinol (Oxf)* 53, 689–695.

McNamara JR & Schaefer EJ. (1987) Automated enzymatic standardized lipid analyses for plasma and lipoprotein fractions. *Clinica chimica acta; Int J Clin Che* 166, 1–8.

Mendelson JH, Sholar MB, Mutschler NH, Jaszyna-Gasior M, Goletiani NV, Siegel AJ & Mello NK. (2003) Effects of intravenous cocaine and cigarette smoking on luteinizing hormone, testosterone, and prolactin in men. *J Pharmacol Exp Therap* 307, 339–348.

Mohr BA, Bhasin S, Link CL, O'Donnell AB & McKinlay JB. (2006) The effect of changes in adiposity on testosterone levels in older men: longitudinal results from the Massachusetts Male Aging Study. *Eur J Endocrinol / Eur Federation Endocrine Soc* 155, 443–452.

Morley JE, Kaiser FE, Perry HM, 3rd, Patrick P, Morley PM, Stauber PM, Vellas B, Baumgartner RN & Garry PJ. (1997) Longitudinal changes in testosterone, luteinizing hormone, and follicle-stimulating hormone in healthy older men. *Metab: Clin Exp* 46, 410–413.

Muller M, den Tonkelaar I, Thijssen JH, Grobbee DE & van der Schouw YT. (2003a) Endogenous sex hormones in men aged 40-80 years. *Eur J Endocrinol / Eur Federation Endocrine Soc* 149, 583–589.

Muller M, van derSchouw YT, Thijssen JH & Grobbee DE. (2003b) Endogenous sex hormones and cardiovascular disease in men. *J Clin Endocrinol Metab* 88, 5076–5086.

Rosner W, Auchus RJ, Azziz R, Sluss PM & Raff H. (2007) Utility, limitations, and pitfalls in measuring testosterone: an endocrine society position statement. *J Clin Endocrinol Metab* 92, 405–413.

Sanders JL, Boudreau RM, Cappola AR, Arnold AM, Robbins J, Cushman M & Newman AB. (2010) Cardiovascular disease is associated with greater incident dehydroepiandrosterone sulfate decline in the oldest old: the cardiovascular health study all stars study. *J Am Geriatr Soc* 58, 421–426.

Schipf S, Haring R, Friedrich N, Nauck MA, Lau K, Alte D, Stang A, Völzke H & Wallaschofski H. (2011) Low total testosterone is associated with increased risk of incident type 2 diabetes mellitus in men: results from the Study of Health in Pomerania (SHIP). *Aging Male* 14, 168–175.

Snyder PJ. (2008) Decreasing testosterone with increasing age: more factors, more questions. *J Clin Endocrinol Metab* 93, 2477–2478.

Svartberg J, Midtby M, Bonaa KH, Sundsfjord J, Joakimsen RM & Jorde R. (2003) The associations of age, lifestyle factors and chronic disease with testosterone in men: the Tromso Study. *Eur J Endocrinol / Eur Federation Endocrine Soc* 149, 145–152.

Svartberg J, von Muhlen D, Mathiesen E, Joakimsen O, Bonaa KH & Stensland-Bugge E. (2006) Low testosterone levels are associated with carotid atherosclerosis in men. *J Intern Med* 259, 576–582.

Torkler S, Wallaschofski H, Baumeister SE, Volzke H, Dorr M, Felix S, Rettig R, Nauck M & Haring R. (2011) Inverse association between total testosterone concentrations, incident hypertension and blood pressure. *Aging Male* 14, 176–182.

Travison TG, Araujo AB, Kupelian V, O'Donnell AB & McKinlay JB. (2007a) The relative contributions of aging, health, and lifestyle factors to serum testosterone decline in men. *J Clin Endocrinol Metab* 92, 549–555.

Travison TG, Araujo AB, O'Donnell AB, Kupelian V & McKinlay JB. (2007b) A population-level decline in serum testosterone levels in American men. *J Clin Endocrinol Metab* 92, 196–202.

Vermeulen A, Kaufman JM & Giagulli VA. (1996) Influence of some biological indexes on sex hormone-binding globulin and androgen levels in aging or obese males. *J Clin Endocrinol Metab* 81, 1821–1826.

Vikan T, Johnsen SH, Schirmer H, Njolstad I & Svartberg J. (2009) Endogenous testosterone and the prospective association with carotid atherosclerosis in men: the Tromso study. *Eur J Epidemiol* 24, 289–295.

Vikan T, Schirmer H, Njolstad I & Svartberg J. (2010) Low testosterone and sex hormone-binding globulin levels and high estradiol levels are independent predictors of type 2 diabetes in men. *Eur J Endocrinol / Eur Federation Endocrine Soc* 162, 747–754.

Wu FC, Tajar A, Pye SR, Silman AJ, Finn JD, O'Neill TW *et al.* (2008) Hypothalamic-pituitary-testicular axis disruptions in older men are differentially linked to age and modifiable risk factors: the European Male Aging Study. *J Clin Endocrinol Metab* 93, 2737–2745.

Yarnell JW, Beswick AD, Sweetnam PM & Riad-Fahmy D. (1993) Endogenous sex hormones and ischemic heart disease in men. The Caerphilly prospective study. *Arterioscler Thromb* 13, 517–520.

Yasui T, Uemura H, Irahara M, Arai M, Kojimahara N, Okabe R, Ishii Y, Tashiro S & Sato H. (2008) Associations of endogenous sex hormones and sex hormone-binding globulin with lipid profiles in aged Japanese men and women. *Clinica chimica acta; Int J Clin Chem* 398, 43–47.

Yeap BB. (2009) Are declining testosterone levels a major risk factor for ill-health in aging men? *Int J Impot Res* 21, 24–36.

Zmuda JM, Cauley JA, Kriska A, Glynn NW, Gutai JP & Kuller LH. (1997) Longitudinal relation between endogenous testosterone and cardiovascular disease risk factors in middle-aged men. A 13-year follow-up of former Multiple Risk Factor Intervention Trial participants. *Am J Epidemiol* 146, 609–617.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Boxplots of the median sex steroid and gonadotropin concentrations at each examination cycle (17th – 21th).

Table S1. Missing outcome and covariate data for the number of observations and participants.

Table S2. Characteristics of the study sample according to the number of missing examinations.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Breast Cancer RESEARCH

RESEARCH ARTICLE Open Access

Reproductive aging-associated common genetic variants and the risk of breast cancer

Chunyan He[1,2*], Daniel I Chasman[3], Jill Dreyfus[4], Shih-Jen Hwang[5], Rikje Ruiter[6], Serena Sanna[7], Julie E Buring[3], Lindsay Fernández-Rhodes[8], Nora Franceschini[8], Susan E Hankinson[9,10], Albert Hofman[6], Kathryn L Lunetta[5,11], Giuseppe Palmieri[12], Eleonora Porcu[7], Fernando Rivadeneira[6,13,14], Lynda M Rose[3], Greta L Splansky[5], Lisette Stolk[13,14], André G Uitterlinden[6,13,14], Stephen J Chanock[15], Laura Crisponi[7], Ellen W Demerath[4], Joanne M Murabito[5,16], Paul M Ridker[3], Bruno H Stricker[6,13] and David J Hunter[9,10,17]

Abstract

Introduction: A younger age at menarche and an older age at menopause are well established risk factors for breast cancer. Recent genome-wide association studies have identified several novel genetic loci associated with these two traits. However, the association between these loci and breast cancer risk is unknown.

Methods: In this study, we investigated 19 and 17 newly identified single nucleotide polymorphisms (SNPs) from the ReproGen Consortium that have been associated with age at menarche and age at natural menopause, respectively, and assessed their associations with breast cancer risk in 6 population-based studies among up to 3,683 breast cancer cases and 34,174 controls in white women of European ancestry. In addition, we used these SNPs to calculate genetic risk scores (GRSs) based on their associations with each trait.

Results: After adjusting for age and potential population stratification, two age at menarche associated SNPs (rs1079866 and rs7821178) and one age at natural menopause associated SNP (rs2517388) were associated with breast cancer risk (p values, 0.003, 0.009 and 0.023, respectively). The odds ratios for breast cancer corresponding to per-risk-allele were 1.14 (95% CI, 1.05 to 1.24), 1.08 (95% CI, 1.02 to 1.15) and 1.10 (95% CI, 1.01 to 1.20), respectively, and were in the direction predicted by their associations with age at menarche or age at natural menopause. These associations did not appear to be attenuated by further controlling for self-reported age at menarche, age at natural menopause, or known breast cancer susceptibility loci. Although we did not observe a statistically significant association between any GRS for reproductive aging and breast cancer risk, the 4th and 5th highest quintiles of the younger age at menarche GRS had odds ratios of 1.14 (95% CI, 1.01 to 1.28) and 1.13 (95% CI, 1.00 to 1.27), respectively, compared to the lowest quintile.

Conclusions: Our study suggests that three genetic variants, independent of their associations with age at menarche or age at natural menopause, were associated with breast cancer risk and may contribute modestly to breast cancer risk prediction; however, the combination of the 19 age at menarche or the 17 age at natural menopause associated SNPs did not appear to be useful for identifying a high risk subgroup for breast cancer.

Introduction

A younger age at menarche and an older age at menopause are well-established risk factors for the development of breast cancer [1]. In the general population, the risk of breast cancer decreases by 10% for each 2-year delay in menarche [2] but increases by 3% for each year that menopause is delayed [3]. These associations are consistent with the hypothesis that breast cancer risk is related to the extent of steroid hormone exposure during a woman's reproductive years, which drives breast mitotic activity and determines the probability of tumorigenic somatic events [4].

Recently, genome-wide association studies (GWAS) have identified several new common genetic loci

* Correspondence: chunhe@iupui.edu
[1]Department of Public Health, Indiana University School of Medicine, 980 West Walnut Street, R3-C241, Indianapolis, IN 46202, USA
Full list of author information is available at the end of the article

associated with either age at menarche or age at natural menopause. Four independent GWAS of age at menarche have identified two novel loci at *LIN28B* and 9q31.2 [5-8], and two GWAS of age at natural menopause have identified four novel loci on chromosomes 5, 6, 19, and 20 [5,9]. Most recently, the ReproGen Consortium, which consisted of these initial GWASs and many additional studies, has conducted expanded meta-analyses for age at menarche [10] and age at natural menopause [11] and reported more new loci identified for each trait. Given the well-established associations of age at menarche and age at natural menopause with breast cancer risk, we set out to assess whether these common genetic loci influence breast cancer risk and whether a genetic risk score (GRS) for these reproductive events might be useful for identifying a high-risk subgroup for breast cancer. Furthermore, since the reproductive risk factors have been observed to be differentially associated with breast cancer by tumor histological subtypes [12-16], we assessed these genetic associations by tumor histological subtypes defined by estrogen receptor (ER) status.

We therefore conducted a meta-analysis of six population-based studies to investigate the association between genetic loci associated with age at menarche or age at natural menopause and breast cancer risk. We assessed 19 and 17 single-nucleotide polymorphisms (SNPs) that have been previously reported to be linked to age at menarche [10] and age at natural menopause [11], respectively, among up to 3,683 breast cancer cases and 34,174 controls in women of European ancestry and evaluated whether these SNPs were differentially associated with breast cancer subtypes defined by ER status in two studies in which such data were available.

Materials and methods

Study population

The ReproGen Consortium was formed by more than 30 studies in the US and Europe to investigate the genetics of reproductive aging traits [10,11]. Our analysis used data from six population-based studies from the ReproGen Consortium: the Nurses' Health Study (NHS), the Women's Genome Health Study (WGHS), the SardiNIA Breast Cancer Study (SardiNIA), the Rotterdam Study I and II (RSI+II), the Framingham Heart Study (FHS), and the Atherosclerosis Risk in Communities Study (ARIC). Each study had at least 200 breast cancer cases. Four studies were prospective cohort studies, one was a nested case-control study, and one was a case-control study. A description of the six studies is provided in Table 1, and more information is given in Additional file 1. Briefly, breast cancer cases occurring in defined populations during specific periods of time were identified by structured questionnaires, medical records, or linkage with a nationwide registry of cancer or death index or both. By the time we conducted this study, the majority of the women in these studies had passed through menopause. As most of the participants in these studies were European whites, we restricted analyses to women of European ancestry. We excluded subjects with missing information on age. Two studies (NHS and WGHS) provided information on the ER status of the breast tumors for a subset of the cases. This information was extracted from medical records. Each study was approved by the relevant local institutional review boards.

Genotype data

We analyzed genotypes for 19 and 17 independent SNPs with reported associations with age at menarche and age at natural menopause, respectively, in the ReproGen Consortium, in which all SNPs achieved genome-wide significance in the meta-analysis of each trait (combined stage 1 and replication P value of less than 1×10^{-8}) [10,11]. None of these SNPs has been reported to be associated with breast cancer risk in previous GWAS and this is likely because of the very stringent P value threshold used to declare genome-wide significance (usually, P values were less than at least 1×10^{-7}). As positive controls, 10 SNPs with consistently reported

Table 1 List of participating studies and number of case and control subjects

Study acronym	Study name	Study design	Case subjects (n = 3,683)	ER^+/ER^- subjects (n = 1,716/371)	Control subjects (n = 34,174)	All subjects (n = 37,857)
NHS-BC	Nurses' Health Study-Breast Cancer	Nested case-control	1,145 (31.1)	807/181	1,142 (3.3)	2,287 (6.0)
WGHS	Women's Genome Health Study	Prospective cohort	1,099 (29.8)	909/190	22,205 (65.0)	23,304 (61.5)
SardBC	SardiNIA Breast Cancer Study	Case-control	809 (22.0)	-	674 (2.0)	1,483 (3.9)
RSI+II	Rotterdam Study I and II	Prospective cohort	216 (5.9)	-	4,261 (12.5)	4,477 (11.8)
FHS	Framingham Heart Study	Prospective cohort	207 (5.6)	-	3,698 (10.8)	3,905 (10.3)
ARIC	Atherosclerosis Risk in Communities Study	Prospective cohort	207 (5.6)	-	2,194 (6.4)	2,401 (6.3)

Data are presented as the number (percentage) of cases, controls, and all subjects. ER, estrogen receptor.

associations with breast cancer as shown in recent GWAS were included [17-19]. All 46 SNPs are listed in Table S1 of Additional file 2. Genotypes used in this analysis have been previously described [10,11]. Complete genotype data from a total of up to 3,683 cases and 34,174 control subjects were available for analysis after the exclusions described in the 'Study population' section.

Breast cancer risk factors

The six studies from the ReproGen Consortium provided information on one or more of the following risk factors for breast cancer: age (continuous, at study entry or diagnosis), age at menarche (continuous, between 9 and 17 years), age at natural menopause (continuous, between 40 and 60 years), age at first live birth (less than 20, 20 to 24, 25 to 29 or no birth, at least 30 years), family history of breast cancer in first-degree relatives (yes/no), alcohol consumption (less than 5, 5 to 15, 15 to 30, at least 30 g/day), parity (0, 1 to 2, at least 3), menopausal hormone therapy (ever/never), oral contraceptive (OC) use (ever/never), and adult body mass index (BMI) (continuous).

Genetic risk score computation

The GRS was calculated on the basis of the 19 and 17 independent SNPs identified in previous studies as being associated with age at menarche and age at natural menopause, respectively [10,11]. As a younger age at menarche and an older age at menopause are independently associated with an elevated breast cancer risk, we computed separate GRSs for a younger age at menarche and an older age at natural menopause. The risk allele was defined as an allele that was associated with a younger age at menarche or an older age at natural menopause. Two methods were used to determine the GRS: a simple count method (count GRS) and a weighted method (weighted GRS). We assumed an additive genetic model for each SNP, applying a linear weighting of 0, 1, or 2 to genotypes containing 0, 1, or 2 risk alleles, respectively. The count method assumes that each SNP contributes equally to the risk of breast cancer. The count GRS was calculated by simply summing the number of risk alleles of each SNP. For the weighted GRS, each SNP was weighted by β-coefficients obtained from the replication studies of recent meta-analyses of two traits [10,11]. The weighted GRS was calculated by multiplying each β-coefficient by the number of corresponding risk alleles (0, 1, or 2) and then summing the products. To simplify interpretation and facilitate comparison with the count GRS, the weighted GRS was further divided by twice the sum of the β-coefficients and then multiplied by the total number of risk alleles. To provide a positive control and also to control for potential confounding by known breast cancer-associated genetic variants, a count GRS was computed on the basis of the 10 SNPs with consistently reported associations with breast cancer [19]: rs2981582, rs3803662, rs11249433, rs7716600, rs13387042, rs889312, rs13281615, rs999737, rs3817198, and rs1045485.

Statistical analysis

In each of the six studies, we performed logistic regression to evaluate the association with breast cancer for each of the 46 candidate SNPs, assuming an additive genetic model. Logistic regression was also used to analyze the association between GRS and breast cancer by including both GRSs for age at menarche and age at natural menopause in the model as the main effects. The GRSs were modeled as continuous variables or categorized into quintiles, and the cutoff points for quintiles were based on the WGHS population, which is the largest prospective cohort population among all participating studies. This approach was applied to each of the six participating studies. Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated from logistic regression. To control for potential confounding by population stratification, we adjusted for the top principal components of genetic variation chosen for each study. We adjusted for age in the main model. To examine whether the genetic association of each of the candidate SNPs or GRSs with breast cancer is mediated through the onset of menarche or natural menopause, we then adjusted for self-reported age at menarche and age at natural menopause in the main model. Other conventional risk factors for breast cancer - including age at first live birth, family history of breast cancer in first-degree relatives, alcohol consumption, parity, menopausal hormone therapy, OC use, and adult BMI - were further included in the model to control for potential confounding in studies which had such data available. To examine whether these genetic associations differ by breast cancer subtypes, in each of the two studies that provided information on ER status, we then investigated the genetic association of each of the candidate SNPs or GRSs with breast cancer in subgroup analysis by ER histological status (positive or negative).

Forest plots were used to present study-specific ORs and 95% CIs. We then performed meta-analyses by using the fixed-effects model to estimate summary ORs from study-specific estimatesthat were weighted by the inverse of the variance of each study. As the meta-analyses restricted to prospective cohort studies or case-control studies yielded similar results, we present results from only the meta-analysis of all six participating studies. We also tested the heterogeneity of associations across studies as well as across different tumor subtypes by using the Q test [20].

All statistical analyses were performed by using SAS version 9.1 software (SAS Institute Inc., Cary, NC, USA). Power calculations were carried out by using Quanto (University of Southern California, Los Angeles, CA, USA). All *P* values were based on two-sided tests and were considered statistically significant if less than 0.05. Because SNPs were selected on the basis of an *a priori* hypothesis, adjustments for multiple comparison tests were not performed.

Results

The six participating studies contributed 3,683 breast cancer cases and 34,174 controls of self-reported white women of European ancestry (Table 1), all with available data on age and the 46 candidate SNPs, and at least one of the conventional risk factors considered. Of the 3,683 cases, about 52% were from the four prospective cohort studies (WGHS, RSI+II, FHS, and ARIC), about 30% were from the nested case-control study in NHS, and about 18% were from the population-based case-control study in SardiNIA. ER status was known for 2,087 cases in the NHS and the WGHS. On average, compared with the controls, the cases had a younger age at menarche and an older age at natural menopause. The expected associations with breast cancer were generally observed for the conventional risk factors across all of the studies (Table S2 of Additional file 3). The associations of the 46 candidate SNPs with age at menarche or age at natural menopause in the six studies were consistent with the original findings from the two meta-analyses [10,11].

Table 2 shows the risk allele frequency and the corresponding per-risk-allele OR of breast cancer for each of the 46 candidate SNPs. The results are arranged in order of the strength of statistical significance (*P* value). The allele frequency for each SNP in the controls was similar to those reported for populations of European descent [21-23]. After adjusting for age and potential population stratification, we found that, among the 19 candidate SNPs for a younger age at menarche, two SNPs, rs1079866 and rs7821178, were significantly associated with breast cancer risk and had corresponding per-risk-allele ORs of 1.14 (96% CI = 1.05 to 1.24; *P* value = 0.003; *P* for heterogeneity = 0.37) and 1.08 (95% CI = 1.02 to 1.15; *P* value = 0.009; *P* for heterogeneity = 0.43), respectively. The SNP rs1079866 is located about 250 kb away from the *INHBA* gene on chromosome 7, whereas SNP rs7821178 is about 181 kb away from the *PXMP3* gene (also known as *PEX3*) on chromosome 8. The strongest GWAS hit for age at menarche, rs7759938 at *LIN28B* on chromosome 6, was not found to be associated with breast cancer risk (*P* value = 0.60). Of the 17 candidate SNPs associated with an older age at natural menopause, one SNP, rs2517388, was significantly associated with breast cancer risk with a per-risk-allele OR of 1.10 (95% CI = 1.01 to 1.20; *P* value = 0.023; *P* for heterogeneity = 0.08). This SNP is an intronic SNP in the *ASH2L* gene on chromosome 8. The study-specific and summary ORs for the three associated SNPs are shown in Figure 1. Further adjustment for conventional risk factors - including age at menarche, age at natural menopause, age at first live birth, family history of breast cancer in first-degree relatives, alcohol consumption, parity, menopausal hormone therapy, OC use, and adult BMI - did not change the results substantially. For candidate loci for age at menarche and age at natural menopause, the findings did not differ materially when we further adjusted for known breast cancer-associated SNPs.

To evaluate the combined effect of candidate SNPs on breast cancer risk, we calculated a GRS for each trait by using either a count GRS or a weighted GRS approach. The mean values of count and weighted GRSs were 20.41 and 20.03, respectively, for age at menarche and 16.21 and 14.32, respectively, for age at natural menopause (Table 3). Based on the count GRS for a younger age at menarche, the OR for breast cancer associated with each point scored, corresponding to 1 risk allele, was 1.01 (95% CI = 1.00 to 1.03) after age and potential population stratification were adjusted for. ORs did not increase linearly across quintiles of GRS for age at menarche (*P* for trend = 0.06). Compared with women in the lowest quintile, women in the fourth and fifth quintiles had ORs for breast cancer of 1.14 (95% CI = 1.01 to 1.28) and 1.13 (95% CI = 1.00 to 1.27), respectively. Results were similar when analyses were performed by using weighted GRS. Overall, we did not observe statistically significant associations between breast cancer risk and age at natural menopause when either count or weighted GRS was used.

In secondary analyses, we then determined whether the associations of the 46 candidate SNPs with breast cancer vary across tumor subtypes defined by ER status in the NHS and the WGHS (Table 4). For the two SNPs (rs1079866 and rs7821178) that had reported associations with age at menarche and that were associated with overall breast cancer risk, we found no statistically significant evidence that the associations differed across subtypes (*P* for heterogeneity = 0.31 and 0.66, respectively), although rs1079866 appeared to have a stronger association with ER^+ tumors (per-allele OR = 1.26; 95% CI = 1.12 to 1.41) than with ER^- tumors (per-allele OR = 1.11; 95% CI = 0.89 to 1.38). Of note, one SNP that had a reported association with age at menarche, rs17188434, had a significantly stronger association with ER^- tumors (per-allele OR = 1.51; 95% CI = 1.15 to 1.98) than with ER^+ tumors (per-allele OR = 1.08; 95% CI = 0.92 to 1.26; *P* for heterogeneity = 0.035). Another SNP that had a reported association with age at

Table 2 Association of candidate single-nucleotide polymorphism loci and the risk of breast cancer

SNP	Gene(s)[a]	Distance from gene	Chromosome	Position (Build 36)	Risk/Reference alleles[b]	Effect allele	Allele frequency[c] of cases/controls[d]	Effect	SE	OR (95% CI)[e]	*P* value[f]	Direction[g]	*P* value for heterogeneity[h]
Age at menarche													
rs1079866	*INHBA*	~250 kb	7	41436618	C/G	c	0.87/0.85	0.132	0.044	1.14 (1.05-1.24)	0.003	++–+-	0.37
rs7821178	*PXMP3*	~181 kb	8	78256392	A/C	a	0.35/0.33	0.081	0.031	1.08 (1.02-1.15)	0.009	++–++	0.43
rs7642134	*VGLL3*	~70 kb	3	86999572	A/G	a	0.39/0.38	0.058	0.03	1.06 (1.00-1.12)	0.056	+++++	0.83
rs10980926	*ZNF483*	Intronic	9	113333455	G/A	g	0.64/0.65	-0.057	0.031	0.94 (0.89-1.00)	0.066	—+–	0.58
rs1398217	*FUSSEL18*	Intronic	18	43006236	G/C	g	0.42/0.42	0.048	0.03	1.05 (0.99-1.11)	0.10	++-+++	0.40
rs17188434	*NR4A2*	~84 kb	2	156805022	C/T	c	0.07/0.07	0.075	0.059	1.08 (0.96-1.21)	0.20	+++—	0.34
rs13187289	*PHF15*	~12 kb	5	133877076	C/G	c	0.80/0.80	-0.038	0.038	0.96 (0.89-1.04)	0.32	—+-+	0.03
rs12617311	*PLCL1*	~195 kb	2	199340810	A/G	a	0.32/0.32	-0.032	0.033	0.97 (0.91-1.03)	0.33	–+-+-	0.81
rs17268785	*CCDC85A*	Intronic	2	56445587	A/G	a	0.84/0.84	0.037	0.041	1.04 (0.96-1.12)	0.37	-+-+++	0.89
rs2002675	*TRA2B, ETV5*	~4 kb, ~135 kb	3	187112262	A/G	a	0.58/0.57	0.022	0.03	1.02 (0.96-1.08)	0.47	++++–	0.92
rs466639	*RXRG*	Intronic	1	163661506	T/C	t	0.12/0.12	-0.032	0.046	0.97 (0.89-1.06)	0.49	-+-+–	0.35
rs1659127	*MKL2*	~28 kb	16	14295806	G/A	g	0.67/0.66	0.022	0.033	1.02 (0.96-1.09)	0.50	+++++-	0.28
rs9635759	*CA10*	~94 kb	17	46968784	G/A	g	0.69/0.70	-0.021	0.034	0.98 (0.92-1.05)	0.54	-+–+-	0.32
rs10899489	*GAB2*	Intronic	11	77773021	C/A	c	0.85/0.84	0.025	0.041	1.03 (0.95-1.11)	0.54	++++–	0.97
rs10423674	*CRTC1*	Intronic	19	18678903	C/A	c	0.67/0.67	0.017	0.032	1.02 (0.96-1.08)	0.59	+++-+-	0.60
rs7759938	*LIN28B*	~26 kb	6	105485647	T/C	t	0.69/0.69	-0.017	0.032	0.98 (0.92-1.05)	0.60	+-+—	0.53
rs2090409	*TMEM38B*	~400 kb	9	108006909	A/C	a	0.33/0.33	0.012	0.031	1.01 (0.95-1.08)	0.69	–++-+	0.21
rs6438424	*3q13.32*	Intergenic	3	119057512	A/C	a	0.50/0.50	0.002	0.029	1.00 (0.95-1.06)	0.94	—+++	0.88
rs6589964	*BSX*	~18 kb	11	122375893	A/C	a	0.48/0.48	-0.001	0.031	1.00 (0.94-1.06)	0.99	–++++	0.20
Age at natural menopause													
rs2517388	*ASH2L*	Intronic	8	38096889	G/T	g	0.17/0.16	0.096	0.042	1.10 (1.01-1.20)	0.023	+++-+-	0.08

Table 2 Association of candidate single-nucleotide polymorphism loci and the risk of breast cancer *(Continued)*

rs4693089	*HEL308*	Intronic	4	84592646	G/A	g	0.47/0.48	-0.06	0.031	0.94 (0.89-1.00)	0.054	———————	0.38
rs11668344	*TMEM150B*	Intronic	19	60525476	A/G	a	0.64/0.64	-0.049	0.03	0.95 (0.90-1.01)	0.11	+———	0.02
rs2307449	*POLG*	Intronic	15	87664932	T/G	t	0.60/0.61	-0.045	0.031	0.96 (0.90-1.02)	0.15	-+–+-	0.20
rs16991615	*MCM8*	Missense	20	5896227	A/G	a	0.07/0.06	0.079	0.06	1.08 (0.96-1.22)	0.18	+++–+	0.80
rs12294104	*C11orf46, PPED2*	~24 kb,~49 kb	11	30339475	T/C	t	0.17/0.17	0.046	0.039	1.05 (0.97-1.13)	0.24	++-++-	0.56
rs12461110	*NLRP11*	Missense	19	61012475	G/A	g	0.64/0.65	-0.036	0.031	0.96 (0.91-1.03)	0.24	–++-+	0.32
rs2303369	*FNDC4*	Intronic	2	27568920	C/T	c	0.62/0.63	-0.031	0.03	0.97 (0.91-1.03)	0.31	–++-+	0.22
rs4886238	*TDRD3*	Intronic	13	60011740	A/G	a	0.35/0.34	0.029	0.031	1.03 (0.97-1.09)	0.35	++——	0.79
rs2277339	*PRIM1*	Missense	12	55432336	T/G	t	0.90/0.89	0.037	0.05	1.04 (0.94-1.14)	0.46	++++–	0.47
rs10852344	*GSPT1, TNFRSF17*	~7 kb, ~42 kb	16	11924420	C/T	c	0.41/0.41	0.020	0.030	1.02 (0.96-1.08)	0.50	++–++	0.87
rs10183486	*TLK1*	Intronic	2	171699217	C/T	c	0.62/0.62	-0.015	0.031	0.99 (0.93-1.05)	0.63	–+–+	0.98
rs4246511	*RHBDL2*	Intronic	1	39152972	T/C	t	0.27/0.27	0.016	0.035	1.02 (0.95-1.09)	0.66	++—+	0.69
rs2153157	*SYCP2L*	Intronic	6	11005474	A/G	a	0.50/0.50	0.011	0.029	1.01 (0.96-1.07)	0.70	++-+-+	0.59
rs1635501	*EXO1*	Intronic	1	240107398	T/C	t	0.53/0.52	0.011	0.034	1.01 (0.95-1.08)	0.74	+-+–+	0.31
rs365132	*UIMC1*	Synonymous	5	176311180	T/G	t	0.50/0.50	-0.009	0.03	0.99 (0.93-1.05)	0.76	+-+–+	0.58
rs1046089	*BAT2*	Missense	6	31710946	G/A	g	0.65/0.65	-0.008	0.031	0.99 (0.93-1.05)	0.79	-+-+-+	0.24
Breast cancer													
rs2981582	*FGFR2*	Intronic	10	123342307	A/G	a	0.44/0.38	0.188	0.03	1.21 (1.14-1.28)	4.7×10^{-10}	++++++	0.10
rs3803662	*TOX3*	~6 kb	16	51143842	A/G	a	0.31/0.25	0.166	0.032	1.18 (1.11-1.26)	2.6×10^{-7}	++++-+	0.31
rs11249433	*FCGR1B*	~245 kb	1	120982136	G/A	g	0.44/0.46	0.149	0.031	1.16 (1.09-1.23)	1.9×10^{-6}	+++++-	0.26
rs7716600	*MRPS30*	~59 kb	5	44910762	A/C	a	0.24/0.20	0.148	0.035	1.16 (1.08-1.24)	2.3×10^{-5}	++++++	0.80
rs13387042	*TNP1*	~181 kb	2	217614077	A/G	a	0.55/0.50	0.114	0.029	1.12 (1.06-1.19)	1.0×10^{-4}	++-+++	0.20
rs889312	*MAP3K1*	~43 kb	5	56067641	C/A	c	0.30/0.32	0.122	0.032	1.13 (1.06-1.20)	1.5×10^{-4}	++++++	0.65

Table 2 Association of candidate single-nucleotide polymorphism loci and the risk of breast cancer *(Continued)*

rs13281615	*8q24.21*	Intergenic	8	128424800	G/A	g	0.43/0.50	0.107	0.03	1.11 (1.05-1.18)	3.0×10^{-4}	++++–	0.01
rs999737	*RAD51L1*	Intronic	14	68104435	C/T	c	0.77/0.80	0.079	0.035	1.08 (1.01-1.16)	0.024	++++-+	0.31
rs3817198	*LSP1*	Intronic	11	1865582	C/T	c	0.32/0.34	0.049	0.032	1.05 (0.99-1.12)	0.13	+++-++	0.45
rs1045485	*CASP8*	Missense	2	201857834	G/C	g	0.88/0.92	0.067	0.045	1.07 (0.98-1.17)	0.13	++–++	0.75

[a]Nearest gene(s); [b]the risk allele refers to the allele associated with a younger age at menarche, an older age at natural menopause, or an increased risk of breast cancer; [c]effect allele frequency; [d]effect allele frequency in case or control subjects, respectively; [e]per effect allele change in log odds ratio (OR) of breast cancer; [f]*P* value from meta-analysis with additive genetic coding after adjustment for age and top genetic principle components; [g]direction of effect allele association with breast cancer in the six studies in order: Nurses' Health Study, Women's Genome Health Study, SardiNIA Breast Cancer Study, Rotterdam Study I and II, Framingham Heart Study, and Atherosclerosis Risk in Communities Study; [h]*P* value from heterogeneity test across studies with 5 degrees of freedom. CI, confidence interval; SE, standard error; SNP, single-nucleotide polymorphism.

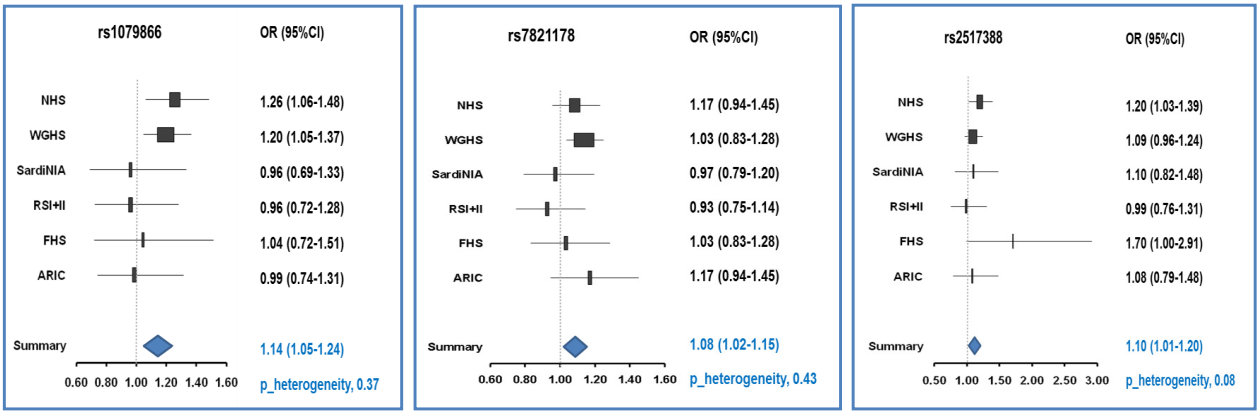


Figure 1 Forest plots for the three candidate loci (rs1079866, rs7821178, and rs2517388) in association with breast cancer risk. Per-risk-allele odds ratios (ORs) and 95% confidence intervals (CIs) were obtained from unconditional logistic regression in each study, and age and potential population stratification were adjusted for. The size of the box is inversely proportional to the standard error of the log OR estimate. *P* values for heterogeneity across studies are 0.37, 0.43, and 0.08, respectively. ARIC, Atherosclerosis Risk in Communities Study; FHS, Framingham Heart Study; NHS, Nurses' Health Study; RSI+II, Rotterdam Study I, II; SardiNIA, SardiNIA Breast Cancer Study; WGHS, Women's Genome Health Study.

Table 3 Association between genetic risk score and risk of breast cancer

		Quintile of GRS					
	Continuous GRS	**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**	***P* for trend**
Count GRS[a]							
Age at menarche							
Cases/Controls	3,683/34,174	677/6,694	739/6,814	731/6,788	783/6,915	753/6,963	
Mean/Median GRS (range)[b]	20.41 (9.12-31.95)	17.00 (9.12-18.09)	19.02 (18.09-19.86)	20.37 (19.86-21.05)	21.90 (21.05-22.74)	23.89 (22.74-31.95)	
OR (95% CI)[c]	1.01 (1.00-1.03)	1.00	1.10 (0.97-1.24)	1.01 (0.89-1.14)	1.14 (1.01-1.28)	1.13 (1.00-1.27)	0.06
Age at natural menopause							
Cases/Controls	3,683/34,174	740/6,699	751/6,948	719/6,787	755/6,929	718/6,811	
Mean/Median GRS (range)[b]	16.21 (4.15-26.52)	12.95 (4.15-14.01)	14.91 (14.01-15.62)	16.15 (15.62-16.90)	17.64 (16.90-18.37)	19.55 (18.37-26.52)	
OR (95% CI)[c]	1.00 (0.98-1.01)	1.00	1.02 (0.90-1.15)	0.93 (0.82-1.05)	1.02 (0.90-1.14)	0.98 (0.87-1.11)	0.54
Breast cancer							
Cases/Controls	3,683/34,174	537/6,864	542/6,525	660/6,716	956/7,352	988/6,717	
Mean/Median GRS (range)[b]	8.88 (1.06-16.96)	6.04 (1.06-7.01)	7.99 (7.01-8.05)	8.99 (8.05-9.07)	9.99 (9.07-10.86)	11.31 (10.86-16.96)	
OR (95% CI)[c]	1.13 (1.11-1.15)	1.00	1.19 (1.04-1.37)	1.38 (1.21-1.57)	1.60 (1.41-1.81)	1.89 (1.67-2.14)	1.5×10^{-31}
Weighted GRS[a]							
Age at menarche							
Cases/Controls	3,683/34,174	762/8,961	711/6,811	752/6,287	728/6,059	736/6,056	
Mean/Median GRS (range)[b]	20.03 (7.15-32.17)	16.12(7.15-17.50)	18.47 (17.50-19.31)	20.07 (19.31-20.83)	21.63 (20.83-22.55)	23.82 (22.55-32.17)	
OR (95% CI)[c]	1.01 (1.00-1.03)	1.00	1.02 (0.90-1.15)	0.99 (0.88-1.13)	1.02 (0.90-1.16)	1.11 (0.98-1.26)	0.12
Age at natural menopause							
Cases/Controls	3,683/34,174	749/7,038	728/6,915	756/6,749	696/6,668	755/6,804	
Mean/Median GRS (range)[b]	14.32 (2.78-26.21)	11.13 (2.78-12.18)	12.96 (12.18-13.61)	14.24 (13.61-14.88)	15.57 (14.88-16.39)	17.63 (16.39-26.21)	
OR (95% CI)[c]	1.00 (0.98-1.01)	1.00	1.07 (0.95-1.20)	0.97 (0.86-1.09)	1.00 (0.88-1.13)	1.03 (0.91-1.16)	0.73

[a]See the 'Materials and methods' section for count genetic risk score (GRS) and weighted GRS computation; [b]mean for continuous GRS and median for each quintile; [c]adjusted for age and potential population stratification. CI, confidence interval; OR, odds ratio.

Table 4 Association of candidate single-nucleotide polymorphism loci and risk of breast cancer by estrogen receptor status in the Nurses' Health Study and Women's Genome Health Study

SNP	All cases		ER^+		ER^-		ER^+/ER^-
	OR (95% CI)[a]	*P* value[b]	OR (95% CI)[a]	*P* value[b]	OR (95% CI)[a]	*P* value[b]	*P* value for heterogeneity[c]
Cases/Controls[d]	2,087/23,319		1,716/23,319		371/23,319		
Age at menarche							
rs1079866	1.22 (1.10-1.35)	2.0×10^{-4}	1.26 (1.12-1.41)	7.3×10^{-5}	1.11 (0.89-1.38)	0.34	0.31
rs7821178	1.12 (1.04-1.20)	3.3×10^{-3}	1.10 (1.01-1.19)	0.025	1.14 (0.98-1.34)	0.10	0.66
rs17188434	1.15 (1.00-1.33)	0.048	1.08 (0.92-1.26)	0.34	1.51 (1.15-1.98)	2.7×10^{-3}	0.035
rs13187289	0.92 (0.84-1.00)	0.055	0.90 (0.82-0.99)	0.026	1.05 (0.87-1.28)	0.60	0.15
rs10980926	0.93 (0.87-1.00)	0.067	0.94 (0.87-1.02)	0.14	0.90 (0.77-1.05)	0.17	0.58
rs7642134	1.06 (0.99-1.14)	0.12	1.07 (0.99-1.16)	0.089	1.09 (0.93-1.27)	0.28	0.86
rs10423674	1.05 (0.97-1.13)	0.22	1.06 (0.97-1.15)	0.19	0.95 (0.81-1.11)	0.50	0.23
rs6589964	0.95 (0.89-1.03)	0.22	0.94 (0.87-1.02)	0.13	1.00 (0.85-1.17)	0.98	0.51
rs1398217	1.04 (0.97-1.12)	0.25	1.02 (0.95-1.10)	0.57	1.11 (0.95-1.29)	0.19	0.36
rs1659127	1.03 (0.96-1.12)	0.41	1.00 (0.92-1.09)	0.98	1.09 (0.92-1.29)	0.32	0.38
rs12617311	0.97 (0.89-1.05)	0.44	1.00 (0.91-1.09)	0.93	0.94 (0.79-1.12)	0.48	0.56
rs17268785	1.04 (0.94-1.14)	0.48	1.07 (0.96-1.19)	0.22	0.83 (0.68-1.00)	0.055	0.023
rs2090409	0.97 (0.90-1.05)	0.50	0.96 (0.89-1.05)	0.38	0.98 (0.84-1.15)	0.82	0.85
rs10899489	1.03 (0.93-1.14)	0.52	1.02 (0.92-1.14)	0.67	0.98 (0.79-1.21)	0.86	0.72
rs6438424	0.98 (0.91-1.05)	0.55	0.99 (0.92-1.07)	0.79	0.98 (0.84-1.14)	0.76	0.88
rs2002675	1.02 (0.95-1.10)	0.55	1.04 (0.96-1.12)	0.35	0.94 (0.81-1.10)	0.45	0.27
rs466639	1.02 (0.92-1.14)	0.73	1.05 (0.94-1.18)	0.39	0.94 (0.74-1.20)	0.62	0.41
rs7759938	1.01 (0.93-1.09)	0.84	1.02 (0.94-1.11)	0.63	1.00 (0.84-1.17)	0.96	0.80
rs9635759	1.01 (0.93-1.09)	0.88	1.00 (0.92-1.09)	0.98	1.03 (0.87-1.22)	0.72	0.74
Age at natural menopause							
rs2517388	1.14 (1.03-1.25)	0.010	1.11 (1.00-1.24)	0.046	1.32 (1.09-1.61)	4.7×10^{-3}	0.12
rs2303369	0.93 (0.86-1.00)	0.036	0.93 (0.86-1.01)	0.076	0.94 (0.80-1.09)	0.42	0.94
rs12461110	0.93 (0.86-1.00)	0.042	0.91 (0.84-0.99)	0.022	0.95 (0.81-1.11)	0.50	0.66
rs4886238	1.06 (0.99-1.14)	0.11	1.06 (0.98-1.15)	0.14	1.03 (0.88-1.21)	0.68	0.76
rs16991615	1.09 (0.94-1.26)	0.23	1.13 (0.97-1.32)	0.12	0.99 (0.71-1.37)	0.94	0.46
rs2277339	1.07 (0.95-1.20)	0.28	1.09 (0.96-1.24)	0.21	1.02 (0.80-1.31)	0.85	0.68
rs12294104	1.04 (0.95-1.14)	0.41	1.06 (0.96-1.17)	0.28	0.98 (0.80-1.21)	0.89	0.54
rs4246511	1.03 (0.95-1.12)	0.42	1.05 (0.96-1.15)	0.29	0.95 (0.79-1.14)	0.59	0.34
rs10852344	1.03 (0.96-1.10)	0.42	1.04 (0.96-1.12)	0.31	0.99 (0.85-1.15)	0.88	0.56
rs2153157	1.03 (0.96-1.10)	0.45	1.04 (0.96-1.12)	0.32	1.00 (0.87-1.17)	0.95	0.69
rs4693089	0.98 (0.91-1.05)	0.54	0.98 (0.91-1.07)	0.68	0.91 (0.77-1.06)	0.23	0.37
rs10183486	0.98 (0.91-1.06)	0.67	0.98 (0.91-1.06)	0.64	0.91 (0.78-1.06)	0.23	0.39
rs2307449	0.99 (0.92-1.06)	0.72	0.96 (0.89-1.04)	0.36	1.06 (0.91-1.24)	0.47	0.29
rs1635501	0.99 (0.91-1.07)	0.76	0.99 (0.90-1.08)	0.80	0.99 (0.83-1.19)	0.94	0.97
rs11668344	1.01 (0.94-1.08)	0.81	1.00 (0.92-1.08)	0.92	1.05 (0.89-1.22)	0.58	0.59
rs1046089	0.99 (0.92-1.07)	0.89	1.02 (0.94-1.10)	0.67	0.90 (0.77-1.05)	0.19	0.17
rs365132	1.00 (0.93-1.07)	0.97	1.00 (0.93-1.08)	1.00	1.00 (0.86-1.16)	0.95	0.96
Breast cancer							
rs11249433	1.20 (1.12-1.29)	4.8×10^{-7}	1.23 (1.14-1.33)	6.9×10^{-8}	1.06 (0.91-1.23)	0.45	0.081
rs3803662	1.20 (1.12-1.30)	1.8×10^{-6}	1.26 (1.16-1.37)	4.2×10^{-8}	0.98 (0.82-1.16)	0.78	0.008
rs2981582	1.18 (1.10-1.27)	8.2×10^{-6}	1.19 (1.10-1.29)	1.1×10^{-5}	1.01 (0.87-1.19)	0.86	0.071
rs13281615	1.16 (1.08-1.24)	3.9×10^{-5}	1.16 (1.07-1.25)	2.1×10^{-4}	1.15 (0.99-1.33)	0.074	0.94
rs13387042	1.15 (1.07-1.23)	8.4×10^{-5}	1.17 (1.08-1.26)	6.3×10^{-5}	1.01 (0.87-1.18)	0.85	0.10
rs7716600	1.17 (1.08-1.27)	1.6×10^{-4}	1.18 (1.08-1.29)	2.9×10^{-4}	1.21 (1.02-1.44)	0.029	0.77
rs889312	1.10 (1.02-1.19)	0.012	1.10 (1.01-1.20)	0.026	1.18 (1.00-1.39)	0.046	0.45
rs999737	1.09 (1.00-1.18)	0.049	1.10 (1.01-1.21)	0.031	0.99 (0.84-1.18)	0.95	0.29

Table 4 Association of candidate single-nucleotide polymorphism loci and risk of breast cancer by estrogen receptor status in the Nurses?'? Health Study and Women?'?s Genome Health Study *(Continued)*

rs1045485	1.09 (0.98-1.20)	0.12	1.06 (0.95-1.19)	0.28	1.20 (0.95-1.52)	0.12	0.35
rs3817198	1.03 (0.96-1.12)	0.38	1.06 (0.98-1.15)	0.16	0.88 (0.74-1.04)	0.13	0.047

[a]Per-risk-allele odds ratio (OR) of breast cancer; [b]*P* value from meta-analysis of the Nurses' Health Study (NHS) and the Women's Genome Health Study (WGHS) with additive genetic coding after adjustment for age and potential population stratification; [c]*P* value from heterogeneity test across estrogen receptor-positive (ER^+) and ER^- tumors; [d]total number of cases or controls in the NHS and the WGHS. CI, confidence interval; SNP, single-nucleotide polymorphism.

menarche, rs17268785, was associated with a decreased risk of ER^- tumors (per-allele OR = 0.83; 95% CI = 0.68 to 1.00) but an increased risk of ER^+ tumors (per-allele OR = 1.07; 95% CI = 0.96 to 1.19; *P* for heterogeneity = 0.023). For the SNP that had a reported association with age at natural menopause and that was associated with overall breast cancer risk, we observed a stronger association with ER^- tumors (per-allele OR = 1.32; 95% CI = 1.09 to 1.61) than ER^+ tumors (per-allele OR = 1.11; 95% CI = 1.00 to 1.24); however, the test for heterogeneity was not statistically significant (*P* for heterogeneity = 0.12). When the count GRS for age at menarche or age at natural menopause was applied to ER^+ and ER^- breast cancer separately, the trend in the OR for ER^+ tumors was very similar to that for overall breast cancer. The ER^- tumor data suggested a somewhat different pattern, although the statistical power was limited for this subtype (Figure 2).

Of the 10 candidate SNPs with consistently reported associations with breast cancer risk, five SNPs (rs11249433, rs3803662, rs2981582, rs13387042, and rs999737) appeared to have a stronger association with ER^+ tumor than ER^- tumors, and rs3803662 reached statistical significance (*P* for heterogeneity = 0.008) with per-risk-allele ORs of 1.26 (95% CI = 1.16 to 1.37) and 0.98 (95% CI = 0.82 to 1.16) for ER^+ and ER^- tumors,

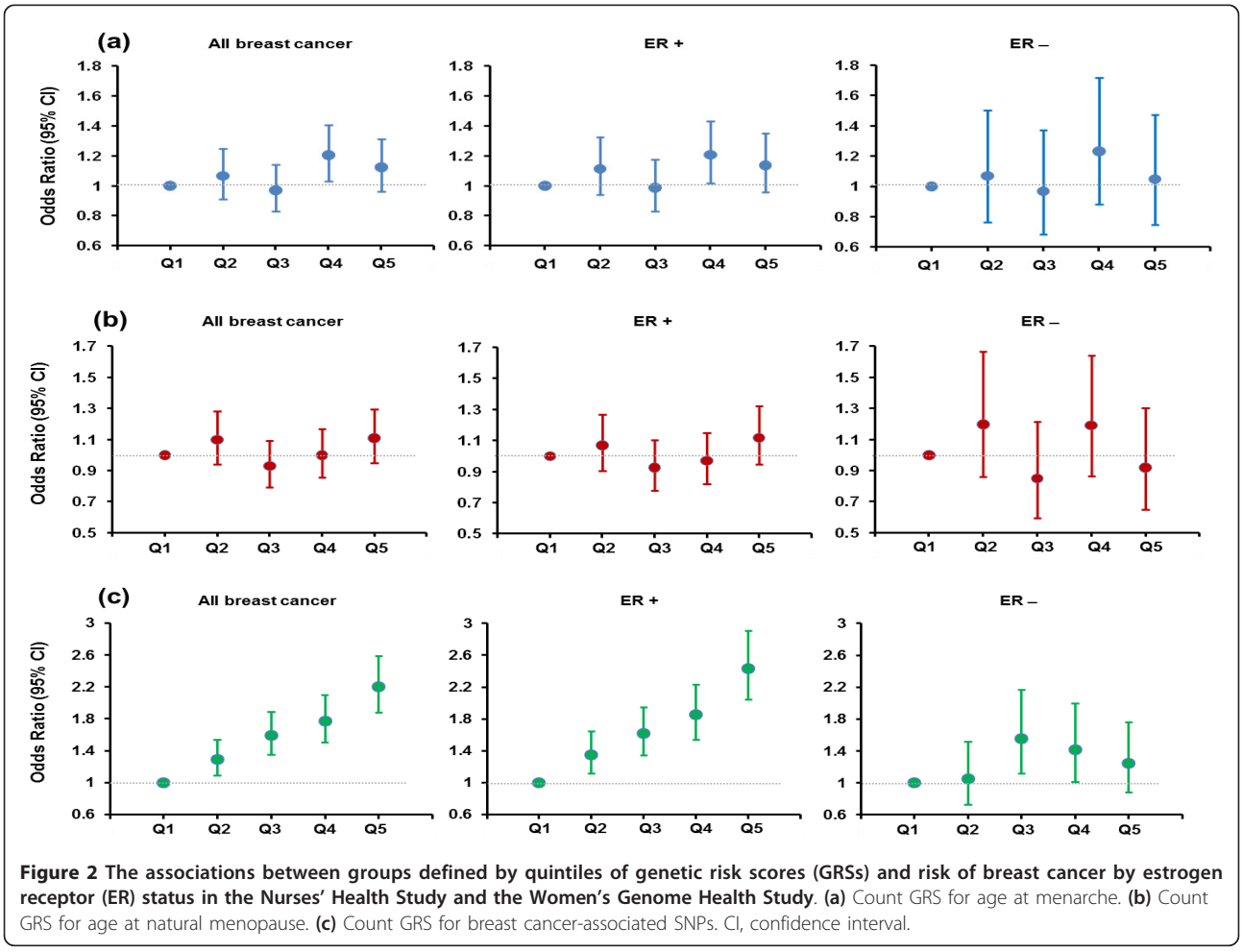


Figure 2 The associations between groups defined by quintiles of genetic risk scores (GRSs) and risk of breast cancer by estrogen receptor (ER) status in the Nurses' Health Study and the Women's Genome Health Study. **(a)** Count GRS for age at menarche. **(b)** Count GRS for age at natural menopause. **(c)** Count GRS for breast cancer-associated SNPs. CI, confidence interval.

respectively (Table 4). Two breast cancer candidate SNPs, rs1045485 and rs3817198, did not show statistically significant associations with overall risk. However, rs1045485 appeared to have a stronger association with ER^- tumors (per-allele OR = 1.20; 95% CI = 0.95 to 1.52) than ER^+ tumors (per-allele OR = 1.06; 95% CI = 0.95 to 1.19; *P* for heterogeneity = 0.35), and rs3817198 was associated with a decreased risk of ER^- tumors (per-allele OR = 0.88; 95% CI = 0.74 to 1.04) but an increased risk of ER^+ tumors (per-allele OR = 1.06; 95% CI = 0.98 to 1.15; *P* for heterogeneity = 0.047).

In these analyses, we further confirmed statistically significant associations with breast cancer risk for 8 of the 10 candidate SNPs that were identified previously in published GWAS of breast cancer (most *P* values were less than 0.001) (Table 2). We did not observe a statistically significant association for either *LSP1*-rs3817198 or *CASP8*-rs1045485 (both with *P* values of 0.13) in our study, although the direction of the associations was consistent with that of previous reports [21,22]. We also calculated, as a positive control, a count GRS based on these 10 SNPs. We found that each score point increase, corresponding to one-risk-allele increase, was significantly associated with an OR of 1.13 (95% CI = 1.11 to 1.15) for breast cancer (Table 3). Compared with women in the lowest quintile, women in the highest quintile had an OR for breast cancer of 1.89 (95% CI = 1.67 to 2.14). For this GRS, the trend in log odds was significantly steeper for ER^+ than for ER^- tumors (*P* for heterogeneity < 0.001), and the OR across quintiles was no longer monotonic in ER^- tumors (Figure 2).

Discussion

In this large meta-analysis of six population-based studies, we investigated whether 19 loci linked with age at menarche and 17 loci linked with age at natural menopause were associated with breast cancer risk among up to 3,683 breast cancer cases and 34,174 controls. We found that two SNPs with reported associations with age at menarche and one SNP with a reported association with age at natural menopause were significantly associated with breast cancer risk. However, no statistically significant associations were found for GRSs that combined all 19 or 17 loci associated with each trait, although the association for age-at-menarche GRS was marginally statistically significant. We confirmed most of the candidate loci for breast cancer which were identified in previous GWAS. Some of these associations appeared to differ by tumor subtypes defined by ER status.

In our analyses, most of the candidate SNPs, including the strongest GWAS hit for age at menarche or age at natural menopause, were not found to be associated with breast cancer risk. This is not necessarily surprising given that age at menarche and age at natural menopause are relatively weak risk factors [2,3], and all candidate SNPs collectively explain only a small portion of the variation of each trait [10,11]. However, two candidate SNPs for age at menarche, rs1079866 and rs7821178, and one candidate SNP for age at natural menopause, rs2517388, were found to be associated with breast cancer risk. These associations were not attenuated after we further adjusted for self-reported age at menarche and age at natural menopause, suggesting these three genetic loci were associated with breast cancer risk independently of their associations with age at menarche or age at natural menopause. It is possible that these genetic loci have pleiotropic effects on reproductive timing as well as other biological processes leading to breast cancer, and the observed associations might be due largely to other biological consequences of these risk variants that do not manifest themselves as changes in age at menarche or age at natural menopause. Alternatively, it is also possible that the relatively crude assignment of these reproductive events to a single chronological year is not sufficiently accurate to capture the biological effect of these processes on breast cancer risk and the genetic variants contribute independent information on the underlying biological risk. The three candidate SNPs also contributed to breast cancer risk independently of the known susceptibility loci for breast cancer, as further adjustment for breast cancer loci did not materially alter the results.

We found statistically significant evidence of association with breast cancer for eight of the 10 breast cancer susceptibility loci examined: *FGFR2*-rs2981582, *TNRC9*-rs3803662, 1p-rs11249433, 5p-rs7716600, 2q35-rs13387042, *MAP3K1*-rs889312, 8q24-rs13281615, and *RAD51L1*- rs999737. The direction and magnitude of these associations were consistent with those of previous reports [17,18,22-25]. We did not observe a statistically significant association for either *LSP1*-rs3817198 or *CASP8*-rs1045485. However, these two SNPs had relatively small reported effects that our study might not have been able to detect. When the 10 candidate SNPs were combined by using a polygenic risk score, the relative risk for women in the highest quintile was about twice that in the lowest quintile, and this is in accordance with other published results [19,26]. In this study, none of the 10 breast cancer susceptibility loci was significantly associated with age at menarche or age at natural menopause, and this is in line with a previous report [27].

Given that most of the candidate loci for age at menarche and age at natural menopause were not associated with breast cancer risk, it is not surprising that there were no statistically significant associations for the polygenetic risk scores that combined all candidate loci

for each trait. To conduct a *post hoc* and exploratory analysis, we created a polygenetic risk score by including only the three candidate loci associated with either age at menarche or age at natural menopause and with breast cancer risk and found that each risk allele increment was associated with an approximately 17% increased risk for breast cancer. Women with four or more risk alleles had an approximately 60% increased risk for breast cancer in comparison with those with two risk alleles or less. When we further combined the three associated SNPs with the 10 breast cancer susceptibility loci to create a polygenetic risk score, each risk allele increment was associated with an approximately 18% increased risk for breast cancer. For women with 14 or more risk alleles (the highest quintile), the risk for breast cancer increased threefold in comparison with those with 10 or less (the lowest quintile). Because the former group constitutes approximately 20% of the study population, the GRS that combines the three candidate SNPs for age at menarche and age at natural menopause and the identified breast cancer susceptibility loci might be useful for identifying a subgroup of women with a high genetic risk for breast cancer. Further research is needed to confirm this finding.

It has been hypothesized that the risk of ER^+ breast cancer is positively associated with a woman's cumulative lifetime exposure to endogenous ovarian hormones [28]. A younger age at menarche [12,15,29] and an older age at menopause [30] have been observed to be more consistently associated with ER^+ than ER^- tumors. In this report, we found that candidate loci for age at menarche and age at natural menopause may also be differentially associated with tumor subtypes defined by ER status. Of the three candidate loci that were found to be associated with overall breast cancer risk, rs1079866 was more strongly associated with ER^+ tumors, rs7821178 was equally associated with both, whereas rs2517388 was more strongly associated with ER^- tumors, although differences were not statistically significant. Importantly, two candidate loci for age at menarche, rs17188434 and rs17268785, had significantly different associations with ER^+ and ER^- tumors. Whereas both SNPs were not significantly associated with overall and ER^+ breast cancer, the former showed a statistically significant positive association with ER^- tumors, whereas the latter showed a statistically significant inverse association with ER^- tumor. These findings provide further support for the notion that ER^+ and ER^- tumors are the result of different etiologic pathways [31].

Although common genetic variants that influence the intermediate phenotypes or risk factors have been hypothesized to subsequently affect disease risk, few studies have assessed the association between these genetic variants and disease risk or, furthermore, whether these associations are mediated through the intermediate phenotypes. Chen and colleagues [32] investigated obesity-linked genetic variants in relation to breast cancer risk but found no statistically significant association. To our knowledge, ours is the first study to evaluate the associations of candidate loci for age at menarche and age at natural menopause with breast cancer risk. One of the strengths of our study is the relatively large combined sample size achieved through international collaboration. We had adequate statistical power (80%) to detect an OR of 1.12 for SNPs with a minor allele frequency (MAF) of 0.10 and an OR of 1.09 for SNPs with an MAF of 0.20. However, our analysis of ER^+ tumors was less adequately powered, as the ER status was not available for all cases, and the study had limited statistical power for ER^- tumors. One limitation in our study is the multiple comparisons that could lead to false-positive results. Although none of the candidate SNPs with a reported association with age at menarche or age at natural menopause survived Bonferroni correction in the test of breast cancer association, this correction is considered to be overly conservative given that the candidates were chosen on the basis of promising hypotheses. Another potential limitation of our study comes from differences in the study population and designs and methods of collecting risk factors and genetic marker data across studies. However, the findings were generally consistent across studies, arguing for the robustness of our results. Finally, as our analyses were restricted to women of European ancestry, results from this study may not be generalizable to other ethnic groups.

Conclusions

In summary, in this large analysis of the association of several novel candidate loci for age at menarche and age at natural menopause with breast cancer risk, we observed that three loci - two for age at menarche and one for age at natural menopause - were significantly associated with breast cancer risk independently of their associations with each trait and independently of known breast cancer susceptibility loci. These associations may differ by tumor subtypes defined by ER status. A combination of all 19 loci associated with age at menarche or 17 loci associated with age at natural menopause did not appear to be helpful for identifying a high-risk subgroup for breast cancer.

Additional material

Additional file 1: Supplementary methods for study population.

Additional file 2: Table S1: Information on the 46 candidate SNP loci identified in previous genome-wide association studies for age at menarche, age at natural menopause and breast cancer.

Additional file 3: Table S2: Characteristics of non-genetic risk factors for breast cancer in each participating study.

Abbreviations

ARIC: Atherosclerosis Risk in Communities Study; BMI: body mass index; CI: confidence interval; ER: estrogen receptor; FHS: Framingham Heart Study; GRS: genetic risk score; GWAS: genome-wide association studies; MAF: minor allele frequency; NHS: Nurses' Health Study; OC: oral contraceptive; OR: odds ratio; RSI+II: Rotterdam Study I and II; SardiNIA: SardiNIA Breast Cancer Study; SNP: single-nucleotide polymorphism; WGHS: Women's Genome Health Study.

Acknowledgements

The NHS breast cancer GWAS scan was performed as part of the Cancer Genetic Markers of Susceptibility initiative of the National Cancer Institute (NCI). We particularly acknowledge the contributions of Robert Hoover, Amy Hutchinson, Kevin Jacobs, and Gilles Thomas. The present research is supported by CA 40356 and U01-CA98233 from the NCI. We acknowledge the study participants in the NHS for their contribution in making this study possible. The WGHS is supported by HL 043851 and HL69757 from the National Heart, Lung, and Blood Institute (NHLBI) and CA 047988 from the NCI, the Donald W Reynolds Foundation, and the Fondation Leducq, and collaborative scientific support and funding for genotyping were provided by Amgen (Thousand Oaks, CA, USA). We thank Joseph Miletich and Alex Parker and the technical staff at Amgen for their collaboration and scientific support in performing the genotyping for the WGHS. The SardiNIA Breast Cancer Study was funded by National Institutes of Health (NIH) contract NO1-AG-1-2109 from the National Institute of Aging (NIA) to the SardiNIA team (principal investigator: Manuela Uda). We thank Grazia Palomba, Antonio Cossu, Francesco Tanda, Mario Budroni, and the other members of the Sardinian Translational Oncology Group for their contribution in collecting Sardinian patients with breast cancer. The generation and management of GWAS genotype data for the Rotterdam Study are supported by The Netherlands Organisation for Scientific Research (NWO) Investments (175.010.2005.011 and 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (RIDE2) (014-93-015) and The Netherlands Genomics Initiative/NWO project 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, and Marjolein Peters for their help in creating the GWAS database and Karol Estrada and Maksim V Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by the Erasmus Medical Center and Erasmus University, Rotterdam; the Netherlands Organisation for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study, and the participating general practitioners and pharmacists. The authors acknowledge the essential role of the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) Consortium in the development and support of this manuscript. CHARGE members consist of The Netherlands' Rotterdam Study, the FHS, the Cardiovascular Health Study, the NHLBI's ARIC study, and the NIA's Iceland Age, Gene/Environment Susceptibility (AGES) study. The FHS phenotype-genotype analyses were supported in part by the NIA (Genetics of Reproductive Life Period and Health Outcomes, R21AG032598; JMM, KLL). The FHS of the NHLBI of the NIH and Boston University School of Medicine was supported by NHLBI's FHS contract N01-HC-25195 and its contract with Affymetrix, Inc. (Santa Clara, CA, USA) for genotyping services (contract N02-HL-6-4278). Analyses reflect intellectual input and resource development from the FHS investigators participating in the SNP Health Association Resource (SHARe) project. A portion of this research was conducted by using the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. The ARIC study is carried out as a collaborative study supported by NHLBI contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367, R01HL086694, and RC2 HL102419; National Human Genome Research Institute contract U01HG004402; and NIH contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by grant UL1RR025005, a component of the NIH, and NIH Roadmap for Medical Research.

Author details

[1]Department of Public Health, Indiana University School of Medicine, 980 West Walnut Street, R3-C241, Indianapolis, IN 46202, USA. [2]Indiana University Melvin and Bren Simon Cancer Center, 535 Barnhill Drive, Indianapolis, IN 46202, USA. [3]Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, 900 Commonwealth Avenue East, Boston, MA 02215, USA. [4]Division of Epidemiology and Community Health, School of Public Health, 1300 South Second Street, Suite 300, University of Minnesota, Minneapolis, MN 55454, USA. [5]Framingham Heart Study, The National Heart Lung and Blood Institute, 73 Mount Wayte, Suite 2, Framingham, MA 01701, USA. [6]Department of Epidemiology, Erasmus Medical Center, 3015 GE, Rotterdam, The Netherlands. [7]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, c/o Cittadella Universitaria di Monserrato, Monserrato, 09042 Cagliari, Italy. [8]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 137 East Franklin Street, Suite 306, CB #8050, Chapel Hill, NC 27599, USA. [9]Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. [10]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, and Harvard medical School, 181 Longwood Avenue, Boston, MA 02115, USA. [11]Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, 3rd floor, Boston, MA 02118, USA. [12]Unit of Cancer Genetics, Istituto di Chimica Biomolecolare, Consiglio Nazionale delle Ricerche, Li Punti, 07100 Sassari, Italy. [13]Department of Internal Medicine, Erasmus Medical Center, 3000 CA, Rotterdam, The Netherlands. [14]Netherlands Consortium of Healthy Aging, Rotterdam, Rotterdam, The Netherlands. [15]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, 6120 Executive Boulevard, Bethesda, MD 20892, USA. [16]Sections of General Internal Medicine, Department of Medicine, Boston University School of Medicine, 720 East Concord Street, Boston, MA 02118, USA. [17]Program in Medical and Population Genetics, Broad Institute of Harvard University and MIT, 301 Binney Street, Cambridge, MA 02142, USA.

Authors' contributions

CH performed the meta-analysis, wrote the manuscript, helped to conceive and design the experiments, helped to perform the primary statistical analyses in each study, shared responsibility for the interpretation of results and critical revision of the manuscript, and contributed to a critical revision of the manuscript for important intellectual content. DIC helped to conceive and design the experiments, helped to perform the primary statistical analyses in each study, shared responsibility for the interpretation of results and critical revision of the manuscript, and contributed to a critical revision of the manuscript for important intellectual content. DJH helped to conceive and design the experiments, contributed to a critical revision of the manuscript for important intellectual content, and participated in the original design, subject recruitment, acquisition of data, biospecimen collection for the studies, and the genotyping and quality control of genotype and other data. JD, SJH, RR, and SS helped to perform the primary statistical analyses in each study and shared responsibility for the interpretation of results and critical revision of the manuscript. SJC, LC, EWD, JMM, PMR, and BHS contributed to a critical revision of the manuscript for important intellectual content and participated in the original design, subject recruitment, acquisition of data, biospecimen collection for the studies, and the genotyping and quality control of genotype and other data. JEB, LF-R, NF, SEH, AH, KLL, GP, EP, FR, LMR, GLS, LS, and AGU participated in the original design, subject recruitment, acquisition of data, biospecimen collection for the studies, and the genotyping and quality control of genotype and other data. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 10 November 2011 Revised: 31 January 2012
Accepted: 20 March 2012 Published: 20 March 2012

References

1. Kelsey JL, Horn-Ross PL: **Breast cancer: magnitude of the problem and descriptive epidemiology.** *Epidemiol Rev* 1993, **15**:7-16.
2. Hsieh CC, Trichopoulos D, Katsouyanni K, Yuasa S: **Age at menarche, age at menopause, height and obesity as risk factors for breast cancer: associations and interactions in an international case-control study.** *Int J Cancer* 1990, **46**:796-800.
3. **Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer.** *Lancet* 1997, **350**:1047-1059.
4. Pike MC, Spicer DV, Dahmoush L, Press MF: **Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk.** *Epidemiol Rev* 1993, **15**:17-35.
5. He C, Kraft P, Chen C, Buring JE, Pare G, Hankinson SE, Chanock SJ, Ridker PM, Hunter DJ, Chasman DI: **Genome-wide association studies identify loci associated with age at menarche and age at natural menopause.** *Nat Genet* 2009, **41**:724-728.
6. Ong KK, Elks CE, Li S, Zhao JH, Luan J, Andersen LB, Bingham SA, Brage S, Smith GD, Ekelund U, Gillson CJ, Glaser B, Golding J, Hardy R, Khaw KT, Kuh D, Luben R, Marcus M, McGeehin MA, Ness AR, Northstone K, Ring SM, Rubin C, Sims MA, Song K, Strachan DP, Vollenweider P, Waeber G, Waterworth DM, Wong A, *et al*: **Genetic variation in LIN28B is associated with the timing of puberty.** *Nat Genet* 2009, **41**:729-733.
7. Perry JR, Stolk L, Franceschini N, Lunetta KL, Zhai G, McArdle PF, Smith AV, Aspelund T, Bandinelli S, Boerwinkle E, Cherkas L, Eiriksdottir G, Estrada K, Ferrucci L, Folsom AR, Garcia M, Gudnason V, Hofman A, Karasik D, Kiel DP, Launer LJ, van Meurs J, Nalls MA, Rivadeneira F, Shuldiner AR, Singleton A, Soranzo N, Tanaka T, Visser JA, Weedon MN, *et al*: **Meta-analysis of genome-wide association data identifies two loci influencing age at menarche.** *Nat Genet* 2009, **41**:648-650.
8. Sulem P, Gudbjartsson DF, Rafnar T, Holm H, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Alexandersen P, Feenstra B, Boyd HA, Aben KK, Verbeek AL, Roeleveld N, Jonasdottir A, Styrkarsdottir U, Steinthorsdottir V, Karason A, Stacey SN, Gudmundsson J, Jakobsdottir M, Thorleifsson G, Hardarson G, Gulcher J, Kong A, Kiemeney LA, Melbye M, Christiansen C, Tryggvadottir L, Thorsteinsdottir U, Stefansson K: **Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche.** *Nat Genet* 2009, **41**:734-738.
9. Stolk L, Zhai G, van Meurs JB, Verbiest MM, Visser JA, Estrada K, Rivadeneira F, Williams FM, Cherkas L, Deloukas P, Soranzo N, de Keyzer JJ, Pop VJ, Lips P, Lebrun CE, van der Schouw YT, Grobbee DE, Witteman J, Hofman A, Pols HA, Laven JS, Spector TD, Uitterlinden AG: **Loci at chromosomes 13, 19 and 20 influence age at natural menopause.** *Nat Genet* 2009, **41**:645-647.
10. Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, Lunetta KL, Visser JA, Byrne EM, Cousminer DL, Gudbjartsson DF, Esko T, Feenstra B, Hottenga JJ, Koller DL, Kutalik Z, Lin P, Mangino M, Marongiu M, McArdle PF, Smith AV, Stolk L, van Wingerden SH, Zhao JH, Albrecht E, Corre T, Ingelsson E, Hayward C, Magnusson PK, Smith EN, *et al*: **Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies.** *Nat Genet* 2010, **42**:1077-1085.
11. Stolk L, Perry JR, Chasman DI, He C, Mangino M, Sulem P, Barbalic M, Broer L, Byrne EM, Ernst F, Esko T, Franceschini N, Gudbjartsson DF, Hottenga JJ, Kraft P, McArdle PF, Porcu E, Shin SY, Smith AV, van Wingerden S, Zhai G, Zhuang WV, Albrecht E, Alizadeh BZ, Aspelund T, Bandinelli S, Lauc LB, Beckmann JS, Boban M, Boerwinkle E, *et al*: **Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways.** *Nat Genet* 2012, **44**:260-268.
12. Althuis MD, Fergenbaum JH, Garcia-Closas M, Brinton LA, Madigan MP, Sherman ME: **Etiology of hormone receptor-defined breast cancer: a systematic review of the literature.** *Cancer Epidemiol Biomarkers Prev* 2004, **13**:1558-1568.
13. Kotsopoulos J, Chen WY, Gates MA, Tworoger SS, Hankinson SE, Rosner BA: **Risk factors for ductal and lobular breast cancer: results from the nurses' health study.** *Breast Cancer Res* 2010, **12**:R106.
14. Li CI, Malone KE, Daling JR, Potter JD, Bernstein L, Marchbanks PA, Strom BL, Simon MS, Press MF, Ursin G, Burkman RT, Folger SG, Norman S, McDonald JA, Spirtas R: **Timing of menarche and first full-term birth in relation to breast cancer risk.** *Am J Epidemiol* 2008, **167**:230-239.
15. Ma H, Bernstein L, Pike MC, Ursin G: **Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies.** *Breast Cancer Res* 2006, **8**:R43.
16. Reeves GK, Pirie K, Green J, Bull D, Beral V: **Reproductive factors and specific histological types of breast cancer: prospective study and meta-analysis.** *Br J Cancer* 2009, **100**:538-544.
17. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, Hankinson SE, Hutchinson A, Wang Z, Yu K, Chatterjee N, Garcia-Closas M, Gonzalez-Bosquet J, Prokunina-Olsson L, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Diver R, Prentice R, Jackson R, Kooperberg C, Chlebowski R, Lissowska J, *et al*: **A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1).** *Nat Genet* 2009, **41**:579-584.
18. Pharoah PD, Antoniou AC, Easton DF, Ponder BA: **Polygenes, risk prediction, and targeted prevention of breast cancer.** *N Engl J Med* 2008, **358**:2796-2803.
19. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, Thun MJ, Cox DG, Hankinson SE, Kraft P, Rosner B, Berg CD, Brinton LA, Lissowska J, Sherman ME, Chlebowski R, Kooperberg C, Jackson RD, Buckman DW, Hui P, Pfeiffer R, Jacobs KB, Thomas GD, Hoover RN, Gail MH, Chanock SJ, Hunter DJ: **Performance of common genetic variants in breast-cancer risk models.** *N Engl J Med* 2010, **362**:986-993.
20. Higgins JP, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** *Stat Med* 2002, **21**:1539-1558.
21. Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, Scollen S, Baynes C, Ponder BA, Chanock S, Lissowska J, Brinton L, Peplonska B, Southey MC, Hopper JL, McCredie MR, Giles GG, Fletcher O, Johnson N, dos Santos Silva I, Gibson L, Bojesen SE, Nordestgaard BG, Axelsson CK, Torres D, Hamann U, Justenhoven C, Brauch H, Chang-Claude J, Kropp S, *et al*: **A common coding variant in CASP8 is associated with breast cancer risk.** *Nat Genet* 2007, **39**:352-358.
22. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, *et al*: **Genome-wide association study identifies novel breast cancer susceptibility loci.** *Nature* 2007, **447**:1087-1093.
23. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, *et al*: **Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2007, **39**:865-869.
24. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover RN, Thomas G, Chanock SJ: **A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.** *Nat Genet* 2007, **39**:870-874.
25. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK, Strobbe LJ, Swinkels DW, van Engelenburg KC, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Saez B, Lambea J, Godino J, Polo E, Tres A, Picelli S, Rantala J, Margolin S, Jonsson T, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, *et al*: **Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2008, **40**:703-706.
26. Reeves GK, Travis RC, Green J, Bull D, Tipper S, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M: **Incidence of breast cancer and its subtypes in relation to individual and multiple low-penetrance genetic susceptibility loci.** *JAMA* 2010, **304**:426-434.

27. Travis RC, Reeves GK, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M: **Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study.** *Lancet* 2010, **375**:2143-2151.
28. Clavel-Chapelon F: **Differential effects of reproductive factors on the risk of pre- and postmenopausal breast cancer. Results from a large cohort of French women.** *Br J Cancer* 2002, **86**:723-727.
29. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL, Gaudet M, Schmidt MK, Broeks A, Cox A, Fasching PA, Hein R, Spurdle AB, Blows F, Driver K, Flesch-Janys D, Heinz J, Sinn P, Vrieling A, Heikkinen T, Aittomaki K, Heikkila P, Blomqvist C, Lissowska J, Peplonska B, Chanock S, Figueroa J, Brinton L, Hall P, Czene K, *et al*: **Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies.** *J Natl Cancer Inst* 2011, **103**:250-263.
30. Phipps AI, Chlebowski RT, Prentice R, McTiernan A, Wactawski-Wende J, Kuller LH, Adams-Campbell LL, Lane D, Stefanick ML, Vitolins M, Kabat GC, Rohan TE, Li CI: **Reproductive history and oral contraceptive use in relation to risk of triple-negative breast cancer.** *J Natl Cancer Inst* 2011, **103**:470-477.
31. Allred DC, Brown P, Medina D: **The origins of estrogen receptor alpha-positive and estrogen receptor alpha-negative human breast cancer.** *Breast Cancer Res* 2004, **6**:240-245.
32. Chen F, Wilkens LR, Monroe KR, Stram DO, Kolonel LN, Henderson BE, Le Marchand L, Haiman CA: **No association of risk variants for diabetes and obesity with breast cancer: the Multiethnic Cohort and PAGE studies.** *Cancer Epidemiol Biomarkers Prev* 2011, **20**:1039-1042.

doi:10.1186/bcr3155
Cite this article as: He *et al.*: Reproductive aging-associated common genetic variants and the risk of breast cancer. *Breast Cancer Research* 2012 **14**:R54.

MAJOR ARTICLE

Risk of Community-Acquired Pneumonia in Veteran Patients to Whom Proton Pump Inhibitors Were Dispensed

John A. Hermos,[1,2] Melissa M. Young,[1] Jennifer R. Fonda,[1] David R. Gagnon,[1,3] Louis D. Fiore,[1,2] and Elizabeth V. Lawler[1,4]

[1]Pharmaco-epidemiology Group, Massachusetts Veterans Epidemiology Research and Information Center, VA Cooperative Studies Program, VA Boston Healthcare System; [2]Department of Medicine, Boston University School of Medicine; [3]Department of Biostatistics, Boston University School of Public Health; and [4]Department of Medicine, Harvard Medical School, Boston, Massachusetts

Background. Observational studies linking proton pump inhibitor (PPI) exposure with community-acquired pneumonia (CAP) have reported either modest or no associations. Accordingly, we studied PPI exposure and CAP in veteran patients, using a retrospective, nested case-control design.

Methods. From linked pharmacy and administrative databases of the New England Veterans Healthcare System, we identified 71 985 outpatients newly prescribed PPIs between 1998 and 2007; 1544 patients met criteria for CAP subsequent to PPI initiation; 15 440 controls were matched through risk-set sampling by age and time under observation. Crude and adjusted odds ratios comparing *current* with *past* PPI exposures, as well as tests for interactions, were conducted for the entire and stratified samples.

Results. Current PPI use associated with CAP (adjusted odds ratio [OR], 1.29 [95% confidence interval {CI}, 1.15–1.45]). Risks were not substantially altered by age or year of diagnosis. Dementia (n = 85; $P = .062$ for interaction) and sedative/tranquilizer use (n = 224; $P = .049$ for interaction) were likely effect modifiers increasing a PPI-CAP association; conversely, for some chronic medical conditions, PPI-associated CAP risks were reversed. PPI exposures between 1 and 15 days increased CAP risks, compared with longer exposures, but PPI initiation also frequently occurred shortly after CAP diagnoses. Prescribed PPI doses >1 dose/day also increased PPI-associated CAP risks.

Conclusions. Among the veterans studied, current compared with past PPI exposures associated modestly with increased risks of CAP. However, our observations that recent treatment initiation and higher PPI doses were associated with greater risks, and the inconsistent PPI-CAP associations between patient subgroups, indicate that further inquiries are needed to separate out coincidental patterns of associations.

Community-acquired pneumonia (CAP) is a common cause of morbidity and mortality, particularly among elderly individuals and those with medical comorbidities [1–3]. Gastric acid suppression is purported to provide a mechanism for increasing risks for CAP by facilitating bacterial colonization of the stomach and upper intestine, potentially leading to colonization of the upper aerodigestive tract with pathogens [4–7]. Among current gastric acid suppressants, proton pump inhibitors (PPIs) provide the most potent and effective management of peptic ulcers, gastroesophageal reflux, and other acid-peptic–related diseases [8, 9]. PPIs have few overt side effects but are often initiated and continued without proven indications [10–13].

Results from retrospective, observational studies within the past decade have been inconclusive regarding the possible association of PPI exposure with CAP [14–19]. Community-based studies from the Netherlands [14], Denmark [15], and Great Britain [16] reported an increased CAP risk of 50%–75% for patients currently prescribed PPIs. Conversely, 3 studies from Great Britain [17] and the United States [18, 19] failed to show such an association. Two meta-analyses of short-term, prospective PPI trials showed no association of PPI use with respiratory infections [20, 21], but recent studies have demonstrated increased risks of recurrent CAP [22] and hospital-acquired pneumonia [23] with

Received 8 April 2011; accepted 13 September 2011; electronically published 18 November 2011.

Correspondence: John A. Hermos, MD, VA Boston Healthcare System (151MAV), 150 S Huntington Ave, Boston, MA 02130 (john.hermos@va.gov).

Clinical Infectious Diseases 2012;54(1):33–42
Published by Oxford University Press on behalf of the Infectious Diseases Society of America 2011.
DOI: 10.1093/cid/cir767

PPI exposures. Thus, evidence linking PPI exposure to CAP and to other respiratory infections remains unsettled.

We have tested the association of PPI use and CAP in a population of US military veterans in a retrospective, nested case-control design. Our objectives were to determine whether current exposure to PPIs increased risks for CAP and whether the duration of current PPI exposure and PPI doses influenced risks. We believed that veteran patients, frequently burdened by chronic medical conditions, would be particularly vulnerable to pneumonia and its consequences and, hence, to this possible associated risk for CAP with PPI use.

METHODS

Setting

We analyzed the linked pharmacy and administrative databases from the New England Veterans Healthcare System (VISN 1) from October 1996 through September 2007. The New England region consists of 8 medical centers and affiliated clinics providing inpatient and outpatient care. VISN 1 pharmacy files were obtained from Information Resource Management (Boston, Massachusetts). Data elements of the outpatient pharmacy files pertinent to our analysis include patient identification, date of birth, drug name and dose, administration route, quantity, date of original prescription, days' supply, refill date(s), and discontinuation date(s). Data on comorbidities were captured by accessing the VA administrative databases (Patient Treatment File and Outpatient Care File) located at the Austin Automation Center (Texas). Analyses were conducted at the Massachusetts Veterans Epidemiology Research and Information Center, VA Boston Healthcare System (Massachuetts). The study was approved by the Institutional Review Board of the VA Boston Healthcare System.

Study Design

We used a retrospective, nested case-control design in which both cases and controls had exposure to a new PPI prescription and from which a definite period of observation was accrued from the first PPI dispense date until the onset of a qualified CAP case or until no VA services were recorded for 18 months. PPI exposures were classified as *current* or *past* in relation to the date of the CAP diagnosis, and comparative risks of current versus past PPI exposures were determined for cases and controls for the entire sample and for stratified samples.

Study Population

Establishment of the cohort is described in Figure 1. Among 103 597 VISN 1 patients receiving a PPI prescription between 1 October 1997 and 30 September 2007, we identified 71 985 meeting criteria for a new outpatient PPI prescription; this required that no PPI prescription had been filled in the VA in the prior calendar year, for established VA patients, or in the first year of recorded VISN 1 treatment, for patients new to VISN 1 after 1 October 1996. These exclusions were done to establish that PPI users who developed CAP and matched controls could be reasonably classified as new users and observed within a measurable period of PPI exposure (see "Exposure Measurement," below). After excluding patients with missing age and those with no follow-up time after their initial prescription, there were 70 042 patients with qualifying new PPI prescriptions. Within this population were identified 2785 cases with coded pneumonia diagnoses and appropriate antibiotic treatment so as to be classified as CAP (see "Outcome Measurement," below), of which 1235 cases preceding the initial PPI prescription were excluded. The resulting number of patients with CAP following PPI initiation was 1544 (6 were excluded because of a gap of ≥18 months following filling of the PPI prescription). Ten controls for each case were selected by risk-set sampling on the basis of age and time under observation from the initial PPI prescription.

Among excluded cases, 147 patients with CAP diagnoses both before and after PPI initiation were not included in the primary analysis as it was felt that many of these patients would represent recurrent CAP cases, making them at risk for PPI-associated recurrent pneumonia [22]. A sensitivity analysis including these cases was conducted to determine whether there was an increase in the overall association of current PPI exposure with CAP.

Outcome Measurement

CAP was established from *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) inpatient and outpatient diagnostic codes 480–482, 485, and 486, which include bacterial pneumonias, specified and unspecified, and exclude specific mentions of viral pneumonia and influenza. To exclude hospital-acquired pneumonias, inpatient pneumonias were required to be primary admitting diagnoses. To be designated as a CAP case, we required that an outpatient antibiotic appropriate for treating CAP was dispensed at hospital discharge or concurrent with the coded outpatient diagnosis. Classes of confirming antibiotics included quinolones, macrolides, β-lactams, and the following specific agents: trimethoprim-sulfamethoxazole, doxycycline, and tetracycline. Whereas these criteria do not include radiologic or culture confirmation [22], we considered them to be relevant evidence of CAP diagnosed and treated in practice and unlikely to bias results either toward or against PPI exposure.

Exposure Measurement

Exposure to a PPI in relation to CAP was categorized as current if the PPI prescription end date occurred after the index case date, or as past if the prescription end date preceded the index case. Sensitivity analysis (performed using an earlier, overlapping sample of CAP cases) allowing for 15-, 30-, and 90-day intervals between the PPI prescription end date and the CAP diagnosis to

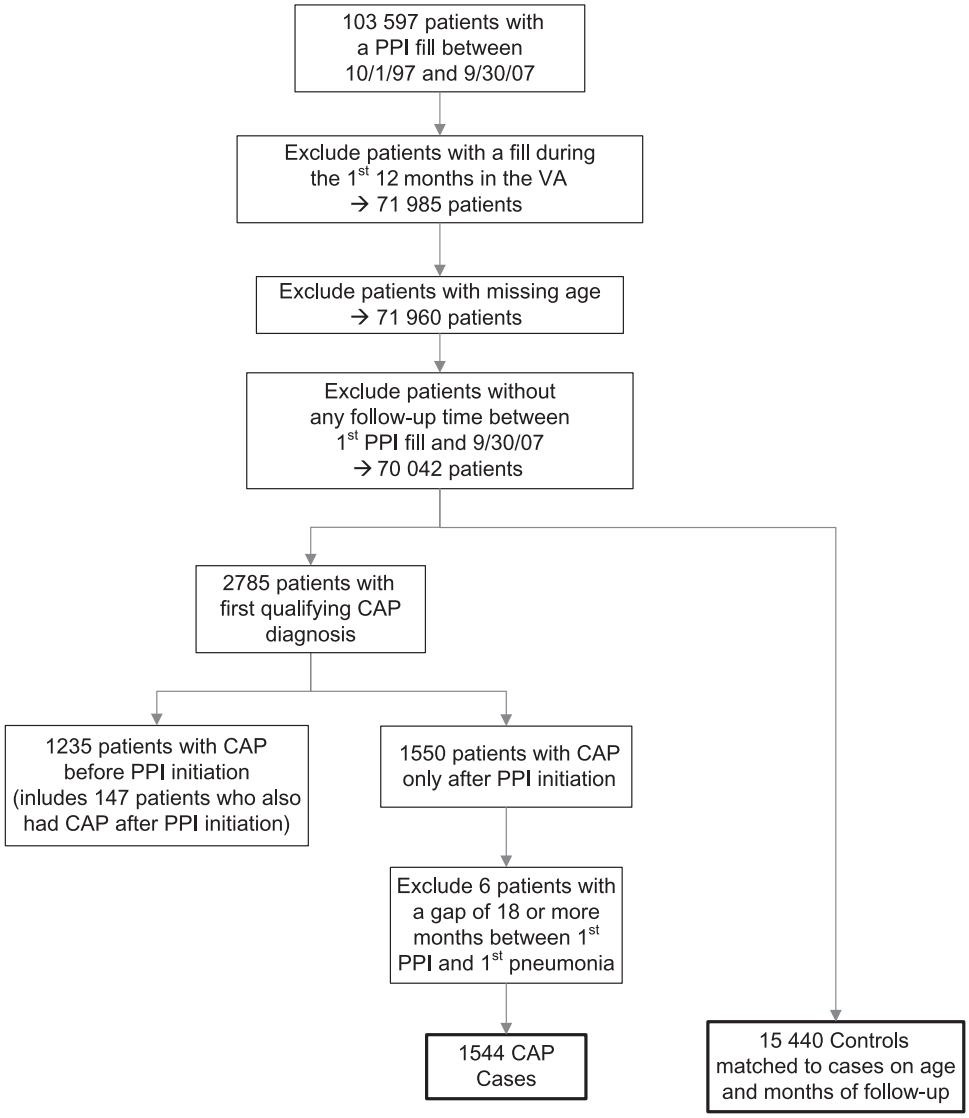


Figure 1. Study population selection. Abbreviations: CAP, community-acquired pneumonia; PPI, proton pump inhibitor.

distinguish current and past use showed no differences or trends in outcomes, so we chose the most conservative criteria. PPI agents in the VA formulary during the observation period and usual prescribed daily doses were omeprazole 20 mg, esomeprazole 20 mg, lansoprazole 15 or 30 mg, pantoprazole 40 mg, and rabeprazole 20 mg. Histamine-2 receptor antagonists (H-2RAs) were not included as a primary exposure because they were available over the counter at reasonable costs so that their use either before or concurrent with PPI use was unknown.

Covariates

Age and comorbidities, including chronic obstructive pulmonary disease (COPD; ICD-9-CM codes 490–494 and 496), congestive heart failure (428), ischemic heart disease (410–414), diabetes (250), lung cancer (162–163.90), nonlung and nonskin malignancies (140–203, excluding 162–163.90 and 173), chronic kidney diseases (585 and 586), cirrhosis (571), dementia (294, 290.40–290.43, 290.0, and 290.10–290.13), alcohol and drug abuse/dependence (303, 304, and 305), gastroesophageal reflux and/or reflux esophagitis (530.81 and 530.1–530.3), and peptic ulcer diseases (531–533), were analyzed as potential confounders as were VA hospital admissions in the 90 days before the index CAP case. Medications active at baseline, H-2RAs, systemic corticosteroids, immunomodulators/immunosuppressants, tranquilizers/sedatives, and antipsychotics, as well as antibiotics prescribed ≤90 days before baseline, were similarly analyzed as covariates.

Data Analysis

Descriptive statistics and odds ratios (ORs) were used to characterize and distinguish cases and controls. Conditional logistic regression was used for all analyses, with all stratified and subgroup analyses performed on the entire matched study population by

Table 1. Baseline Characteristics of Patients With Community-Acquired Pneumonia and Controls

Characteristic	Cases (n = 1544)	Controls (n = 15 440)	OR (95% CI)	*P* Value
Age, mean years (SD)	65.8 (12.2)	65.8 (12.2)	1.00 (.97–1.03)	.978
Male sex	1491 (96.6)	14 755 (95.6)	1.32 (.99–1.75)	.062
Comorbidities diagnosed on or before baseline				
Gastroesophageal reflux	675 (43.7)	6882 (44.6)	0.97 (.87–1.07)	.517
Peptic ulcer	193 (12.5)	1521 (9.9)	1.31 (1.12–1.54)	.001
H. pylori infection	<11 (0.65)	103 (0.67)	0.97 (0.51–1.86)	.929
Lung cancer	77 (5.0)	195 (1.3)	4.09 (3.12–5.35)	<.001
Nonskin, nonlung cancer	346 (22.4)	2301 (14.9)	1.68 (1.47–1.91)	<.001
Ischemic heart disease	711 (46.1)	5827 (37.7)	1.45 (1.30–1.62)	<.001
Chronic kidney disease	91 (5.9)	567 (3.7)	1.65 (1.31–2.08)	<.001
Chronic liver disease	71 (4.6)	432 (2.8)	1.70 (1.31–2.21)	<.001
Dementia	85 (5.5)	541 (3.5)	1.61 (1.27–2.04)	<.001
Alcohol/drug dependence or abuse	624 (40.4)	4092 (26.5)	2.10 (1.86–2.36)	<.001
No. of above diagnoses				
1	391 (25.3)	5368 (34.8)	1.82 (1.49–2.21)	<.001
2	373 (24.2)	3646 (23.6)	2.47 (2.02–3.02)	<.001
≥3	613 (39.7)	2552 (16.5)	5.87 (4.79–7.20)	<.001
None	167 (10.8)	3874 (25.1)	…	
Admission ≤90 days before end date	304 (19.7)	709 (4.6)	5.21 (4.49–6.05)	<.001
Medications active at baseline				
H2-receptor antagonist	383 (24.8)	3188 (20.7)	1.27 (1.12–1.44)	.001
Systemic corticosteroid	152 (9.8)	499 (3.2)	3.25 (2.69–3.93)	<.001
Immunomodulator/immunosuppressant	11 (0.71)	54 (0.35)	2.04 (1.07–3.90)	.032
Tranquilizer/sedative	224 (14.5)	1680 (10.9)	1.39 (1.20–1.62)	<.001
Antipsychotic	124 (8.0)	689 (4.5)	1.91 (1.56–2.34)	<.001
Antibiotic (≤90 days before baseline)	409 (26.5)	2178 (14.1)	2.21 (1.95–2.49)	<.001

Data are no. (%) of patients, unless otherwise indicated.

Abbreviations: CI, confidence interval; *H. pylori*, *Helicobacter pylori*; OR, odds ratio; SD, standard deviation.

requesting specific contrasts in PROC LOGISTIC. We used past PPI users as the reference groups among both cases and controls. Specific stratifying variables included age group (<65 years and ≥65 years), to examine the role of Medicare eligibility, and year of CAP diagnosis (fiscal years 1998–2003 and 2004–2007), to account for potential PPI exposures from nonprescription formulations purchased over the counter during recent years. Finally, we used conditional logistic regression on stratified samples to evaluate statistical interaction for covariates. These analyses were in some cases limited by small stratified samples and were not necessarily expected to reach statistical significance at $P < .05$, but they could show potential trends as effect modifiers.

We calculated unadjusted and adjusted ORs comparing durations of current PPI exposures before the index case. To further describe the time course of PPI use in relation to new CAP diagnoses, we provide frequency distributions among CAP patients that describe the timing of PPI initiation. To conduct this analysis, we included all available CAP cases, including those occurring before PPI initiation (1235 cases were excluded; Figure 1) and those occurring subsequent to PPI initiation (1544 cases). We also compared risks between 196 CAP cases receiving >1 standard PPI dose per day to 434 cases receiving 1 standard PPI dose per day. This analysis excluded 238 patients receiving lansoprazole, as 2 dose formulations, 15 mg and 30 mg, were dispensed in varying amounts, precluding our establishing "standard" daily doses.

All analyses were performed using SAS software, version 9.2 (SAS Institute). We used *P* values <.05 and 95% confidence intervals (CIs) to test for statistical significance.

RESULTS

Patient Characteristics

The study population of 1544 cases and 15 440 controls was approximately 96% male, with a mean age (standard deviation) of 65.8 (12.2) years. (Table 1). With the exception of esophageal reflux diagnoses, cases were more likely than controls to have each and >1 coded medical comorbidity, to have been hospitalized ≤90 days before the index case date, and to be prescribed each of the designated medications.

Table 2. Unadjusted and Adjusted Odds Ratios Associating Proton Pump Inhibitor Exposure and Community-Acquired Pneumonia

Population	PPI Exposure	Cases, No.	Controls, No.	Unadjusted OR (95% CI)	Adjusted OR[a] (95% CI)
All	Current	868	7743	1.29 (1.16–1.43)	1.29 (1.15–1.45)
	Past	676	7697	Reference	Reference
Age ≥65 years	Current	498	4608	1.19 (1.04–1.37)	1.21 (1.04–1.41)
	Past	383	4202	Reference	Reference
Age <65 years	Current	370	3135	1.42 (1.21–1.68)	1.41 (1.19–1.68)
	Past	293	3495	Reference	Reference
FY 1998–FY 2003	Current	328	2436	1.38 (1.15–1.66)	1.39 (1.14–1.69)
	Past	218	2195	Reference	Reference
FY 2004–FY 2007	Current	540	5307	1.24 (1.09–1.42)	1.24 (1.08–1.43)
	Past	458	5502	Reference	Reference

Abbreviations: CI, confidence interval; FY, fiscal year, OR, odds ratio; PPI, proton pump inhibitor.

[a] Adjusted for sex; age; FY of end date (1998–2003 or 2004–2007); baseline *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnoses of chronic obstructive pulmonary disease, diabetes, congestive heart failure, lung cancer, nonskin and nonlung cancers, ischemic heart disease, chronic kidney disease, chronic liver disease, dementia, alcohol/drug dependence or abuse, peptic ulcer, and reflux/esophagitis; admission ≤90 days before end date; active use at baseline of H2-receptor antagonist, systemic corticosteroid, immunomodulator/immunosuppressant, tranquilizer/sedative, and antipsychotic agents; and antibiotic use ≤90 days before baseline.

Patient Population Pneumonia Risks

For the total population of cases, the crude and adjusted odds of CAP among those currently using PPIs were modestly elevated relative to the odds among those with past PPI use (adjusted OR, 1.29 [95% CI, 1.15–1.45]), with no substantial differences in crude and adjusted risks when stratified by age or year of CAP diagnosis (Table 2). When the 147 cases who had CAP both before and subsequent to initiation of PPIs were included in the analysis, the risks were modestly elevated (adjusted OR, 1.40 [95% CI, 1.26–1.56]), indicating an increased PPI-CAP association with recurrent cases.

Confounding and Effect Modification by Comorbidities and Medications

To determine potential confounding and effect modification, adjusted risks and tests for interaction were conducted for comorbidities and medications (Table 3). Cases receiving immunosuppressants showed the highest relative risks for current PPI use, but the number of cases was too small (n = 13) to demonstrate statistical significance. For patients with dementia (n = 85; $P = .062$ for interaction), use of sedatives/tranquilizers (n = 224; $P = .049$ for interaction), and, to a lesser extent, patients with chronic liver disease (n = 71, $P = .257$ for interaction) and use of antipsychotics (n = 124; $P = .258$ for interaction), there was possible effect modification favoring a PPI-CAP association. Of interest, confounding by congestive heart failure, COPD, lung cancer, and chronic kidney disease diagnoses, and having ≥3 comorbidities, showed reduced adjusted risks of CAP with current PPI exposures (ie, a potential protective effect), with congestive heart failure demonstrating effect modification ($P = .014$ for interaction). There was no evidence of confounding or effect modification for esophageal reflux and peptic ulcer diagnoses; for diagnoses of other chronic medical conditions, specifically diabetes, ischemic heart disease, and nonlung and nonskin cancers; and for use of other medications, specifically H-2RAs, corticosteroids, and antibiotics (data not shown).

Duration of Exposure

We determined the ORs for increasing durations of current PPI exposure for predicting pneumonia, using a >180-day exposure as the reference group and including past exposure in the analysis model (Table 4). Those in the exposure group of 1–15 days demonstrated increased risks, compared with those with longer exposures; this short-term exposure group represented only 6.9% of the 868 cases who were exposed to PPIs at the time of their pneumonia. The remaining larger group of patients with longer durations of current PPI exposure also had increased risks for CAP, compared with those patients with past PPI exposures. We determined the frequency distribution of CAP cases occurring both before (n = 1235) and after (n = 1544) PPI initiation, and it showed a time-dependent distributions of cases (Figure 2). Among 330 CAP cases occurring within 90 days before PPI initiation, half occurred within 13 days, which is further indication that CAP cases were likely to occur closely before, as well as closely after, the time of PPI initiation.

Dose Effect

CAP cases were more likely to be currently prescribed PPIs at doses >1 standard dose per day than at doses ≤1 standard dose per day (adjusted OR, 1.33 [95% CI, 1.06–1.65]; $P = .012$) (see Methods for exclusion of 238 patients prescribed lansoprazole at 15-mg or 30-mg doses) (Table 5).

DISCUSSION

From a nested cohort of US veteran patients who predominantly were male, we found a 30% increased risk of CAP with current

Table 3. Odds Ratios and Tests for Effect Modification in the Association Between Proton Pump Inhibitor Exposure and Community-Acquired Pneumonia

Comorbidity/Medication		PPI Exposure	Cases, No.	Controls, No.	Unadjusted OR (95% CI)	Adjusted OR[a] (95% CI)	*P* Value for Interaction[b]
Reflux/esophagitis	Yes	Current	420	3777	1.36 (1.16–1.61)	1.43 (1.21–1.70)	.405
		Past	255	3105	Reference	Reference	
	No	Current	448	3966	1.24 (1.08–1.43)	1.19 (1.03–1.39)	
		Past	421	4592	Reference	Reference	
Peptic ulcer	Yes	Current	105	788	1.11 (.83–1.51)	1.00 (.72–1.37)	.328
		Past	88	733	Reference	Reference	
	No	Current	763	6955	1.31 (1.17–1.47)	1.34 (1.18–1.51)	
		Past	588	6964	Reference	Reference	
Dementia	Yes	Current	51	232	2.00 (1.26–3.18)	2.05 (1.24–3.38)	.062
		Past	34	309	Reference	Reference	
	No	Current	817	7511	1.26 (1.13–1.41)	1.26 (1.12–1.41)	
		Past	642	7388	Reference	Reference	
Alcohol or drug abuse/dependence	Yes	Current	341	1913	1.37 (1.16–1.63)	1.29 (1.08–1.55)	.490
		Past	283	2179	Reference	Reference	
	No	Current	527	5830	1.28 (1.11–1.46)	1.29 (1.12–1.49)	
		Past	393	5518	Reference	Reference	
Lung cancer	Yes	Current	44	106	1.13 (.66–1.92)	0.93 (.52–1.66)	.637
		Past	33	89	Reference	Reference	
	No	Current	824	7637	1.29 (1.16–1.44)	1.31 (1.16–1.47)	
		Past	643	7608	Reference	Reference	
COPD	Yes	Current	382	1613	1.22 (1.03–1.44)	1.14 (.96–1.37)	.420
		Past	301	1527	Reference	Reference	
	No	Current	486	6130	1.32 (1.15–1.52)	1.39 (1.21–1.61)	
		Past	375	6170	Reference	Reference	
Congestive heart failure	Yes	Current	151	700	0.96 (.74–1.24)	0.87 (.66–1.15)	.014
		Past	134	587		Reference	
	No	Current	717	7043	1.36 (1.21–1.53)	1.39 (1.23–1.57)	
		Past	542	7110		Reference	
Chronic liver disease	Yes	Current	43	205	1.72 (1.03–2.88)	1.80 (1.04–3.12)	.257
		Past	28	227	Reference		
	No	Current	825	7538	1.27 (1.14–1.42)	1.27 (1.13–1.43)	
		Past	648	7470	Reference	Reference	
Chronic kidney disease	Yes	Current	52	313	1.09 (.70–1.71)	1.02 (.63–1.64)	.476
		Past	39	254	Reference	Reference	
	No	Current	816	7430	1.30 (1.16–1.45)	1.31 (1.16–1.47)	
		Past	637	7443	Reference	Reference	
≥3 Comorbidities[c]	Yes	Current	337	1318	1.17 (.98–1.40)	1.07 (.88–1.29)	.151
		Past	276	1234	Reference	Reference	
<3 Comorbidities[c]	No	Current	531	6425	1.38 (1.20–1.58)	1.42 (1.24–1.63)	
		Past	400	6463	Reference	Reference	
≥1 VA admissions ≤90 days before end date	Yes	Current	191	389	1.44 (1.09–1.91)	1.36 (1.01–1.83)	.328
		Past	113	320	Reference	Reference	
	No	Current	677	7354	1.24 (1.10–1.39)	1.28 (1.13–1.44)	
		Past	563	7377	Reference	Reference	
Tranquilizer/sedative	Yes	Current	141	846	1.68 (1.26–2.24)	1.77 (1.31–2.40)	.049
		Past	83	834	Reference	Reference	
	No	Current	727	6897	1.23 (1.10–1.38)	1.22 (1.08–1.38)	
		Past	593	6863	Reference	Reference	

Table 3 continued.

Comorbidity/Medication		PPI Exposure	Cases, No.	Controls, No.	Unadjusted OR (95% CI)	Adjusted OR[a] (95% CI)	*P* Value for Interaction[b]
Antipsychotic	Yes	Current	71	315	1.60 (1.09–2.35)	1.76 (1.17–2.66)	.258
		Past	53	374	Reference	Reference	
	No	Current	797	7428	1.27 (1.14–1.42)	1.26 (1.12–1.41)	
		Past	623	7323	Reference	Reference	
Immunosuppressant	Yes	Current	(<11)	31	1.99 (.48–8.35)	2.64 (.58–12.03)	.549
		Past	(<11)	23	Reference	Reference	
	No	Current	860	7712	1.28 (1.15–1.43)	1.28 (1.15–1.44)	
		Past	673	7674	Reference	Reference	

Abbreviations: CI, confidence interval; COPD, chronic obstructive pulmonary disease; OR, odds ratio; PPI, proton pump inhibitor.

[a] Adjusted for sex; age; fiscal year of end date (1998–2003 or 2004–2007); baseline *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnoses of chronic obstructive pulmonary disease, diabetes, congestive heart failure, lung cancer, nonskin and nonlung cancers, ischemic heart disease, chronic kidney disease, chronic liver disease, dementia, and alcohol/drug dependence or abuse; admission ≤90 days before end date; active use at baseline of H2-receptor antagonist, systemic corticosteroid, immunomodulator/immunosuppressant, tranquilizer/sedative, and antipsychotic agents; and antibiotic use ≤90 days before baseline.

[b] Model includes age, sex, given covariate, and interaction term.

[c] Excludes esophageal reflux and peptic ulcer diseases.

compared with past PPI exposures; these findings are between those from prior observational studies that showed positive [14–16] and no [17–19] associations. Three large, community-based, case-controlled studies from the Netherlands (adjusted OR, 1.73 [95% CI, 1.33–2.25]) [14], Denmark (adjusted OR, 1.5 [95% CI, 1.3–1.7]) [15], and Great Britain (adjusted OR, 1.55 [95% CI, 1.36–1.77]) [16] associated current PPI exposure with CAP. Conversely, 3 recently published studies from Great Britain [17] and the United States [18, 19] reported no statistically significant associations between PPI exposures and CAP. In 2 of these latter studies, only patients aged ≥65 years were included [18, 19]. The 3 studies with negative findings included as potential confounders current smoking status [17–19], past year hospitalizations and office visits [17, 19], and a measure of functional status [19], variables potentially attenuating any independent risks conveyed by PPI use. We used common comorbidities, medications, and recent hospitalizations as potential confounders, but we recognize that other key confounders may not have been included.

There was apparent effect modification for patients with dementia and for those receiving sedatives/tranquilizers, each independently increasing this association, perhaps because of impaired swallowing and increased reflux of acid-suppressed stomach contents [24–26]. We also found potential confounding with chronic liver disease and immunosuppressant use, but tests for interactions were limited by small sample sizes. We cannot readily explain our findings that, for other chronic medical conditions, the association between current PPI exposure and CAP was "protective," but it is possible that the inherent risks for CAP with these conditions preclude any added risks incurred by

Table 4. Duration of Current Proton Pump Inhibitor Exposure and Community-Acquired Pneumonia

	Cases		Controls			
PPI duration[a]	No.	Follow-up, Mean Months (SD)	No.	Follow-up, Mean Months (SD)	Unadjusted OR (95% CI)	Adjusted OR[b] (95% CI)
1–15 days	60	16.5 (23.7)	354	30.8 (24.6)	1.52 (1.14–2.04)[c]	1.25 (.91–1.71)[c]
16–45 days	81	26.0 (29.3)	830	22.6 (24.8)	0.88 (.69–1.12)	0.74 (.56.96)
46–90 days	133	25.1 (25.6)	1207	20.1 (24.4)	0.99 (.81–1.22)	0.88 (.71–1.09)
91–180 days	142	23.8 (24.4)	1288	24.0 (25.9)	0.99 (.81–1.21)	0.88 (.72–1.09)
≥181 days	452	36.5 (25.9)	4064	35.1 (24.9)	Reference	Reference
Past users	676	32.6 (26.1)	7697	34.4 (26.1)	0.79 (.70–.90)[c]	0.77 (.67–.88)[c]

Abbreviations: CI, confidence interval; OR, odds ratio; PPI, proton pump inhibitor; SD, standard deviation.

[a] Exposure calculated from end date back, stopping at a gap in treatment of ≥30 d.

[b] Adjusted for sex; age; fiscal year of end date (1998–2003 or 2004–2007); baseline *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnoses of chronic obstructive pulmonary disease, diabetes, congestive heart failure, lung cancer, nonskin and nonlung cancers, ischemic heart disease, chronic kidney disease, chronic liver disease, dementia, and alcohol/drug dependence or abuse; admission ≤90 days before end date; active use at baseline of H2-receptor antagonist, systemic corticosteroid, immunomodulator/immunosuppressant, tranquilizer/sedative, and antipsychotic agents; and antibiotic use ≤90 days before baseline.

[c] $P < .05$.

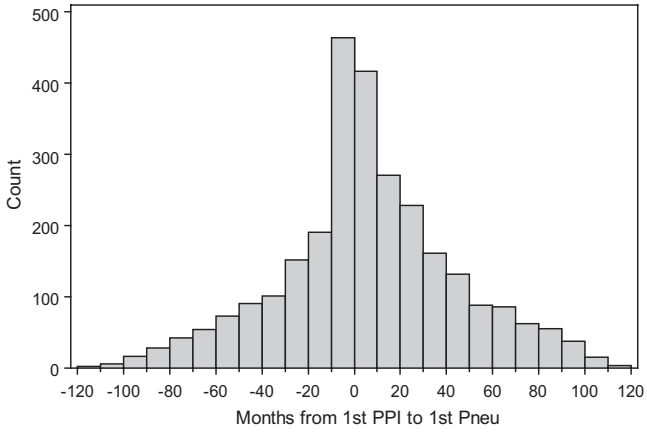


Figure 2. Interval between proton pump inhibitor initiation and community-acquired pneumonia (CAP) for included and excluded CAP cases. Abbreviations: Pneu, pneumonia; PPI, proton pump inhibitor.

PPI exposures [1–3]. These seemingly contradictory findings with respect to specific comorbidities, in this study and others, weaken the argument for a convincing PPI-CAP association in patient populations in which multiple confounders, entered into or omitted from analyses, could alter results.

We, and others [14, 15, 17], found that short-duration current PPI use increased the association with CAP, compared with longer-duration current use; in the other prior observational studies, dose effects were not tested [16, 18, 19]. This finding raises the question as to whether the association results from PPIs being newly prescribed for either early, nonspecific symptoms of pneumonia, for gastric protection, or for new-onset or worsening reflux, or whether it simply occurs because patients are in medical care settings where PPI are commonly prescribed. Our time-interval data from CAP cases that preceded PPI initiation suggests that this may well be the case, as PPI initiations frequently follow CAP diagnoses in the context of clinical care of symptomatic patients. However, >90% of current PPI exposures in our sample were >15 days and also carried modestly increased risks for CAP, compared with risks among past users.

Two prior studies [14, 17] and ours demonstrated a PPI dose–dependent association with CAP. Whereas this finding indicates enhanced, more-potent gastric acid suppression, it is possible that increased PPI dosing is in response to symptomatic gastroesophageal reflux, which might in itself increase risks for chronic aspiration and lower respiratory tract infections [27, 28]. Because we found that PPI prescribing often followed CAP diagnoses, it is likely that some clinicians initiated high-dose PPIs in this clinical context.

Our findings must be viewed with caution because our exposure and outcome measures are subject to limitations inherent in database dependent observational studies. We were not able to determine PPI compliance or supplementation with other PPIs or gastric acid suppressants; however, because of the low costs of VA copayments, it is unlikely that many patients would use non–VA-prescribed prescriptions. We could not identify the indication for PPI use explicitly, although we collected an exhaustive list of conditions potentially related to PPI use or CAP development. We did not validate our CAP cases with documentation of radiographic findings, which is typically necessary for randomized trials [29]. Dublin et al [19], using only coded CAP diagnoses, reported a 30% increased risk for CAP with PPI exposure, but this association was negated using only chart-reviewed, validated CAP cases. However, Gulmez et al [15] found no difference in increased PPI-associated risks in a comparison of patients with and patients without positive chest radiographic findings. We required timely prescribing of antibiotics appropriate for treatment of CAP to establish our CAP cases, but this practice, although certainly indicative of clinically relevant CAP, may considerably vary between physicians and institutions. Even without radiography or other means for case validation, it seems unlikely that misclassifications of CAP cases would bias results toward patients receiving PPIs.

Despite these limitations, our study has considerable strengths. The veteran patient population studied is large, and our 8-year window of observation is long. The VA pharmacy database allows for an accurate determination of physician-entered

Table 5. Association Between Proton Pump Inhibitor Dose at End Date and Community-Acquired Pneumonia

PPI Dose	Cases, No. (%)	Controls, No. (%)	Unadjusted OR (95% CI)	*P* Value	Adjusted OR[a] (95% CI)	*P* Value
>1 standard dose[b]	196 (31.1)	1498 (23.4)	1.42 (1.16–1.73)	<.0001	1.33 (1.06–1.65)	.012
≤1 standard dose[b]	434 (68.9)	4902 (76.6)	Reference		Reference	

Data exclude patients receiving lansoprazole as PPI at end date (see text).

Abbreviations: CI, confidence interval; OR, odds ratio; PPI, proton pump inhibitor.

[a] Adjusted for sex; age; fiscal year of end date (1998–2003 or 2004–2007); baseline *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnoses of chronic obstructive pulmonary disease, diabetes, congestive heart failure, lung cancer, nonskin and nonlung cancers, ischemic heart disease, chronic kidney disease, chronic liver disease, dementia, and alcohol/drug dependence or abuse; admission ≤90 d before end date; active use at baseline of H2-receptor antagonist, systemic corticosteroid, immunomodulator/immunosuppressant, tranquilizer/sedative, and antipsychotic agents; and antibiotic use ≤90 d before baseline.

[b] Standard dose is defined as 1 pill/day of basic dose of each PPI formulation (see Methods for exposure variables).

prescriptions that are actually dispensed. Our study design is conservative in that we used time under observation, used risk set sampling for our overall and stratified analyses, and excluded known recurrent CAP cases. By identifying cases and selecting matched controls only from PPI users, cases and controls likely had reasonably similar indications for initial PPI prescribing, although these could not be identified explicitly. As PPIs are commonly prescribed and continued for protracted periods without clear indications [9–13], even a small or uncertain added risk for CAP could translate into many CAP cases. Along with the risks reported from observational studies for other infections, specifically *Clostridium difficile* infection [30–32], other enteric infections [33], spontaneous bacterial peritonitis [34], recurrent CAP, and hospital-acquired pneumonia [23], consideration of reasons for starting or continuing PPIs or for alternative gastric acid suppression may be warranted [10–13]. Prospective, randomized trials to test these associations would necessarily be prolonged and costly. Thus, additional observational studies, with well-considered potential confounders and targeted, stratified subsamples, that can further test for the existence and extent of associations of PPI exposure with serious infections are needed.

Notes

Acknowledgments. The authors acknowledge the programming support of Galina Sokoloskaya.

Financial support. This work was supported by Core Funds Massachusetts Veterans Epidemiology Research and Information Center, VA Cooperative Studies Program, VA Boston Healthcare System.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Garibaldi RA. Epidemiology of community-acquired respiratory tract infections in adults: incidence, etiology, and impact. Amer J Med **1985**; 78:32–7.
2. Bartlett JG, Mundy LM. Community-acquired pneumonia. N Engl J Med **1995**; 333:1618–24.
3. Fry AM, Shay DK, Holman RC, et al. Trends in hospitalization for pneumonia among persons aged 65 and older in the United States, 1988–2002. JAMA **2005**; 294:2712–19.
4. Giannella RA, Broitman SA, Zamcheck N. Influence of gastric acidity on bacterial and parasitic enteric infections: a perspective. Ann Intern Med **1973**; 78:271–6.
5. Thorens J, Froehlich F, Schwizer W, et al. Bacterial overgrowth during treatment with omeprazole compared with cimetidine: a prospective randomized double blind study. Gut **1996**; 39:54–9.
6. Brummer RJ, Stockburgger RW. Effect of nizatidine 300 mg at night and omeprazole 20 mg in the morning on 24-hour intragastric pH and bacterial overgrowth in patients with acute duodenal ulcer. Dig Dis Sci **1996**; 41:2048–54.
7. Theisen J, Nehra D, Citron D, et al. Suppression of gastric acid secretion in patients with gastroesophageal reflux disease results in gastric bacterial overgrowth and deconjugation of bile acids. J Gastrointest Surg **2000**; 4:50–4.
8. Moss SF, Armstrong D, Arnold R, et al. GERD 2003: a consensus on the way ahead. Digestion **2003**; 67:111–7.
9. Katz PO, Zavala S. Proton pump inhibitors in the management of GERD. J Gastrointest Surg **2010**; 14(Suppl 1):S62–6.
10. Jacobson BC, Ferris TG, Shea TL, Mahlis EM, Lee TH, Wang TC. Who is using chronic acid suppression therapy and why? Am J Gastroenterol **2003**; 98:51–8.
11. Gregor JC. Acid suppression and pneumonia: a clinical indication for rational prescribing. JAMA **2004**; 292:2012–3.
12. Heidelbaugh JJ, Goldberg KL, Inadomi JM. Overutilization of proton pump inhibitors: a review of cost-effectiveness and risk. Am J Gastroenterol **2009**; 104(Suppl 2):S27–32.
13. Yang Y, Metz DC. Reviews in basic and clinical gastroenterology and hepatology: safety of proton pump inhibitor exposure. Gastroenterology **2010**; 139:1115–27.
14. Laheij RJ, Sturkenboom MC, Hassing RJ, Deileman J, Stricker BH, Jansen JB. Risk of community-acquired pneumonia and use of gastric acid-suppressive drugs. JAMA **2004**; 292:1955–60.
15. Gulmez SE, Holm A, Frederiksen H, Jensen TG, Pedersen C, Hallas J. Use of proton pump inhibitors and the risk of community-acquired pneumonia. Arch Intern Med **2007**; 167:950–5.
16. Myles PR, Hubbard RB, McKeever TM, Pogson Z, Smith CJ, Gibson JE. Risk of community-acquired pneumonia and the use of statins, ace inhibitors and gastric acid suppressants: a population-based case-control study. Pharmacoepidemiol Drug Saf **2009**; 18:269–75.
17. Sarkar M, Hennessey S, Yang Y. Proton pump inhibitor use and the risk of community-acquired pneumonia. Ann Intern Med **2008**; 149:391–9.
18. Gau G, Acharya U, Khan S, Heh V, Mody L, Kao T. Pharmacotherapy and the risk for community- acquired pneumonia. BMC Geriatr **2010**; 10:45–52.
19. Dublin S, Walker RL, Jackson ML, Nelson JC, Weiss NS, Jackson LA. Use of proton pump inhibitors and H2 blockers and risk of pneumonia in older adults: a population-based case-control study. Pharmacoepidemiol Drug Saf **2010**; 19:792–802.
20. Sultan N, Nazareno J, Gregor J. Association between proton pump inhibitors and respiratory infections: a systematic review and meta-analysis of clinical trials. Can J Gastroenterol **2008**; 22:761–6.
21. Estborn L, Joelson S. Occurrence of community-acquired respiratory tract infection in patients receiving esomeprazole: retrospective analysis of adverse events in 31 clinical trials. Drug Saf **2008**; 31:627–36.
22. Herzig SJ, Howell MD, Ngo LH, Marcantinio ER. Acid-suppressive medication use and the risk for hospital-acquired pneumonia. JAMA **2009**; 301:2120–8.
23. Loeb M. Pneumonia in older persons. Clin Infect Dis **2003**; 37:1335–9.
24. Malone DC, McLaughlin TP, Wahl PM, et al. Burden of Alzheimer's Disease and association with negative health outcomes. Am J Manag Care **2009**; 15:481–8.
25. Trifiro G, Gambassi G, Sen EF, et al. Association of community-acquired pneumonia with antipsychotic drug use in elderly patients. Ann Intern Med **2010**; 152:418–25.
26. Hogan WJ. Spectrum of supra-esophageal complications of gastroesophageal reflux disease. Amer J Med **1997**; 103:77S–83S.
27. Mays EE, Dubois JJ, Hamilton GB. Pulmonary fibrosis associated with tracheobronchial aspiration: a study of the frequency of hiatal hernia and gastroesophageal reflux in interstitial pulmonary fibrosis of obscure etiology. Chest **1976**; 69:512–5.
28. Niedermann MS, Mandell LA, Anzueto A, et al. American Thoracic Society Guidelines for the management of adults with community-acquired pneumonia. Am J Respir Crit Care Med **2001**; 163: 1730–54.
29. Dial S, Delaney JA, Barkun AN, Suissa S. Use of gastric acid-suppressive agents and the risk of community-acquired *Clostridium difficile*-associated disease. JAMA **2005**; 294:2989–95.
30. Lowe DO, Mamdani MM, Kopp A, Low DE. Proton pump inhibitor use and risk of community-acquired *Clostridium difficile*-associated disease: a population based study. Clin Infect Dis **2006**; 43:1272–6.

31. Linsky A, Gupta K, Lawler EV, Fonda JR, Hermos JA. Proton pump inhibitors and the risk for recurrent *Clostridium difficile* infection. Arch Intern Med **2010**; 170:772–8.
32. Leonard J, Marshall JK, Moayyedi P. Systematic review of the risk of enteric infection in patients taking acid suppression. Am J Gastroenterol **2007**; 102:2047–56.
33. Bajaj JS, Zadvornova Y, Heuman DM, et al. Association of proton pump inhibitors therapy with spontaneous bacterial peritonitis in cirrhotic patients with ascites. Am J Gastroenterol **2009**; 104:1130–4.
34. Campbell MS, Obstein K, Reddy KR, Yang YX. Association between proton pump inhibitor use and spontaneous bacterial peritonitis. Dig Dis Sci **2008**; 53:394–8.

Journal of the American College of Cardiology
© 2012 by the American College of Cardiology Foundation
Published by Elsevier Inc.

Vol. 60, No. 7, 2012
ISSN 0735-1097/$36.00
http://dx.doi.org/10.1016/j.jacc.2011.12.057

SPECIAL ARTICLE

Future Directions for Cardiovascular Disease Comparative Effectiveness Research

Report of a Workshop Sponsored by the National Heart, Lung, and Blood Institute

Mark A. Hlatky, MD,* Pamela S. Douglas, MD,† Nakela L. Cook, MD, MPH,‡ Barbara Wells, PHD,‡ Emelia J. Benjamin, MD, SCD,§ Kay Dickersin, PHD, MA,‖ David C. Goff, MD, PHD,¶ Alan T. Hirsch, MD,# Elaine M. Hylek, MD,§ Eric D. Peterson, MD, MPH,† Véronique L. Roger, MD, MPH,** Joseph V. Selby, MD, MPH,†† James E. Udelson, MD,‡‡ Michael S. Lauer, MD‡

Stanford and Oakland, California; Durham and Winston-Salem, North Carolina; Bethesda and Baltimore, Maryland; Boston, Massachusetts; and Minneapolis and Rochester, Minnesota

Comparative effectiveness research (CER) aims to provide decision makers with the evidence needed to evaluate the benefits and harms of alternative clinical management strategies. CER has become a national priority, with considerable new research funding allocated. Cardiovascular disease is a priority area for CER. This workshop report provides an overview of CER methods, with an emphasis on practical clinical trials and observational treatment comparisons. The report also details recommendations to the National Heart, Lung, and Blood Institute for a new framework for evidence development to foster cardiovascular CER, and specific studies to address 8 clinical issues identified by the Institute of Medicine as high priorities for cardiovascular CER.
(J Am Coll Cardiol 2012;60:569–80)

Comparative effectiveness research (CER) has recently emerged as a national priority, spurred by healthcare reform and economic stimulus legislation. Congress appropriated $1.1 billion for CER as part of the American Recovery and Reinvestment Act of 2009 and is anticipated to enable additional annual spending of $500 million as part of the newly established Patient-Centered Outcomes Research Institute. According to congressional legislation, the Patient-Centered Outcomes Research Institute will give priority for project management to the National Institutes of Health and the Agency for Health Research and Quality. Therefore, the National Heart, Lung, and Blood Institute (NHLBI) may have new opportunities to advance CER related to cardiovascular disease, which remains the leading cause of death and disability in the United States today.

The NHLBI sponsors workshops to solicit input and recommendations on important topics, so on July 13 and 14, 2010, the Division of Cardiovascular Sciences convened a workshop on CER in cardiovascular disease. The workshop brought together 25 outside experts from a variety of disciplines (clinical trials, epidemiology, biostatistics, health services research, and clinical medicine) to discuss a range of future opportunities that NHLBI could consider in CER as it relates to cardiovascular disease and the specific priorities for CER in cardiovascular disease that were identified by the Institute of Medicine (IOM). The discussions at the workshop, therefore, represent the opinions and recommendations of the participants and are not necessarily the policy or priorities of NHLBI.

From the *Stanford University School of Medicine, Stanford, California; †Duke University Medical Center, Durham, North Carolina; ‡National Heart, Lung, and Blood Institute, Bethesda, Maryland; §Boston University School of Medicine, Boston, Massachusetts; ‖Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; ¶Wake Forest University Health Sciences, Winston-Salem, North Carolina; #University of Minnesota Medical School, Minneapolis, Minnesota; and the **Mayo Clinic, Rochester, Minnesota.††Division of Research, Kaiser Permanente of Northern California, Oakland, California; and the ‡‡Tufts University School for Medicine, Boston, Massachusetts. Drs. Cook, Wells, and Lauer are full-time employees of the National Heart, Lung, and Blood Institute (NHLBI). The views expressed in this manuscript do not necessarily represent the views of the NHLBI, National Institutes of Health, or any other government entity. Dr. Goff is a research consultant for and a data and safety monitoring board member for a clinical trial of a glucose-lowering medication marketed by Merck and Takeda. Dr. Hirsch has relationships with Pozen Medical, Merck, AstraZeneca, Abbott Vascular, Viromed, Summit Doppler, Novartis, and Cytokinetics. Dr. Hylek is on the scientific advisory board of Bayer, Boehringer-Ingelheim, Bristol-Myers Squibb, Daiichi Sankyo, Johnson & Johnson, Merck, Ortho-McNeil, and Pfizer. Dr. Selby was employed by Kaiser Permanente and is currently the Executive Director of the Patient Centered Outcomes Research Institute (PCORI); the opinions expressed in this paper do not necessarily reflect those of PCORI. Dr. Udelson has received research funding from the NHLBI. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose. Drs. Hlatky and Douglas cochaired the workshop.

Manuscript received November 14, 2011; revised manuscript received December 16, 2011, accepted December 20, 2011.

Abbreviations and Acronyms

CER = comparative effectiveness research
CT = computed tomography
IOM = Institute of Medicine
NHLBI = National Heart, Lung, and Blood Institute
RCT = randomized controlled trial(s)

This report summarizes the deliberations and recommendations to the NHLBI of this workshop: "Future Directions for Cardiovascular Disease Comparative Effectiveness Research." The report is divided into several sections, including: 1) an overview of CER data sources and methods; 2) a proposed framework for CER at the NHLBI (Fig. 1); and 3) possible approaches to 8 priority CER topics identified by the IOM in the areas of cardiovascular and peripheral vascular disease.

Overview of CER

Comparative effectiveness research has been defined as "the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers and policy makers in making informed decisions that will improve healthcare at both the individual and population levels" (1). CER can focus on care at the patient level or the system level, but regardless of its scope, CER is intended to provide information that can facilitate medical decision making and improve health outcomes (2).

There is great public value in rigorous studies comparing alternative strategies for diagnosis and treatment, as demonstrated by many landmark NHLBI-sponsored clinical trials (3–13). Nevertheless, comparative effectiveness studies that would facilitate healthcare decisions are not performed as often as they are needed, and gaps persist between the production of scientific evidence and the needs of consumers and healthcare providers for evidence on CER. Because there are limited resources to support biomedical research, it is necessary to prioritize key clinical questions that can be answered with comparative effectiveness studies, while expanding as much as possible the pool of investigators capable of performing CER.

Study Designs for CER

A variety of methods are used in CER, including randomized trials, observational studies, simulations and models, systematic reviews, meta-analyses, and collaborative pooling of individual patient data from multiple studies. Randomized trials can be used to compare a management strategy (preventative, diagnostic, therapeutic) with the best alternative strategy. Analyses of data from clinical registries, electronic health records, and administrative databases can address CER questions regarding situations in which randomization may be difficult. Whereas observational studies using existing data may be simpler and less expensive to conduct, they are more susceptible to bias introduced by selection of patients for alternative treatments. Statistical approaches to analysis of observational data can be used to

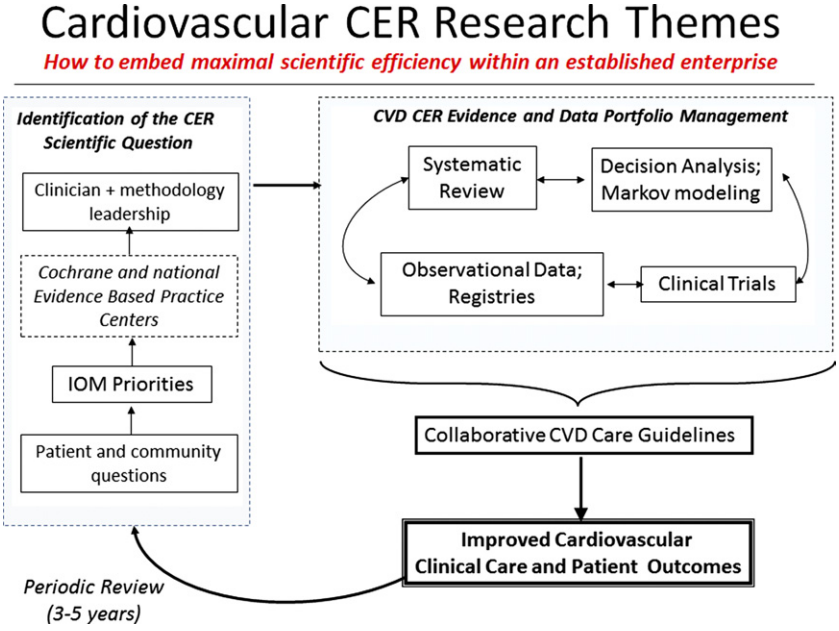


Figure 1 Conceptual Framework for CER

CER = comparative effectiveness research; CVD = cardiovascular disease; IOM = Institute of Medicine.

minimize these biases, but these methods need to be further developed. Decision models and simulations may also be very useful methods in CER. The workshop participants explored the unique opportunities, strengths, and limitations of these methods for use in CER, as summarized herein.

Randomized trials. Randomized clinical trials (RCT) may be applied in CER to compare treatments (e.g., use of a drug, device, procedure, or a behavioral intervention), clinical evaluation strategies (e.g., biomarkers, imaging), healthcare delivery methods (e.g., disease management programs, specialist vs. generalist care), and policy interventions (e.g., copayments, formulary restrictions, regionalization of procedures). The strengths of the RCT include its use of a prospective protocol, standardized data collection, complete follow-up, and random treatment assignment, each of which enhances its "internal validity." An RCT is particularly appropriate when a high level of evidence is required to change practice, such as when optimal management is controversial, the effect of the treatment on outcomes is modest, or the intervention is costly. Although RCT often involve highly selected patients and atypical practice settings, which can limit generalizability, RCT can be designed to enroll more representative patient populations and can be performed in more typical practice settings.

There are many practical barriers to conducting large, representative clinical trials in the United States. The focus on delivering care efficiently and uniformly may not readily accommodate clinical research in routine clinical practice. Building the infrastructure to perform multicenter clinical trials can be difficult, time-consuming, and costly. Thus, exploring strategies to promote more efficient clinical trials is important, such as adopting "large practical trial designs" or fostering the conduct of a range of CER trials by existing investigator or site teams. Use of real-world settings, such as private practice groups, community health centers, and integrated healthcare systems, may enhance recruitment of representative patients in CER trials.

Monetary barriers can impede the conduct of clinical trials, because the cost of research data collection, clinic visits and tests, and follow-up is high. Thus, identifying ways to streamline data collection while maintaining accuracy and validity over the length of follow-up would enhance the ability to conduct CER trials. Health insurers may balk at covering the costs of clinic visits and tests for patients in a clinical trial, which increases the cost to the research budget and may lead patients to drop out of the study. Furthermore, in some instances, financial incentives in healthcare may not be aligned for randomized trials; for example, trials that randomize patients to procedures compared with medical therapy may result in foregone procedural fees. The duration of follow-up in clinical trials is also constrained by cost, which limits collection of long-term data on comparative efficacy and safety.

Barriers for CER trials also include lack of time in busy practices to enroll participants in research studies and low levels of academic recognition for site investigators who are just 1 member of a large trial team. Whereas institutional review boards are needed to protect participants in clinical trials, the multiplicity of jurisdictions across multisite trials complicates the conduct of the trial. Centralized or simplified institutional review processes for multicenter studies should be explored, particularly for comparative trials of existing, approved interventions.

Clinical registries. A clinical registry is the prospective collection of standardized data on the clinical characteristics and outcomes of patients defined by a particular disease, diagnosis, procedure, or exposure. Clinical registries share some features with randomized trials, such as standardized data collection, but unlike clinical trials, they do not dictate patient treatment by either random assignment or a strict protocol. Clinical registries can have broader inclusion criteria than clinical trials do, and, therefore, they may better represent the diversity of patients, providers, and practice settings found in contemporary clinical care. Clinical registries occupy a middle ground between the formal structure of a randomized trial and the collection of relatively unstructured data from medical charts, electronic health records, or claims data.

Clinical registries have evolved from small case series to national (or international) collaborations that enroll thousands to millions of patients. Standardized data definitions and data collection methods are key features of a high-quality clinical registry, as they ensure comparable levels of details about each patient enrolled. Documenting outcomes—particularly outcomes that occur late after a single episode of care—is also essential if clinical registries are to be used in CER. Longer follow-up can be obtained actively by contacting patients (as is done in clinical trials) or passively by linking registry data to electronic health records, claims data, or state and national mortality files.

Clinical registries have been used to define contemporary practice patterns, document disparities in care, and assess the safety of cardiovascular drugs, devices, and procedures in clinical practice. Clinical registries can also be analyzed to compare alternative treatments, but these studies require advanced biostatistical methods to reduce the biases introduced by nonrandomized patient selection for treatment.

Clinical registries may also be linked with randomized clinical trials, as when a registry prospectively collects data on patients screened for entry into a trial or on patients eligible for a trial who decline to be randomized (14). Randomized trials can also capitalize on an ongoing clinical registry, which can be used to identify eligible patients and capture clinical data. Hybrid registry trial designs may be particularly powerful tools for CER, as they permit efficient patient enrollment and provide information of the generalizability of trial results.

The resources required to establish and maintain clinical registries have been provided by a variety of mechanisms. Professional societies have sponsored notable clinical registries often supported by hospital-paid fees for participation,

such as the American Heart Association's Get with the Guidelines Programs (in cardiovascular disease, heart failure, and stroke), the Society of Thoracic Surgeon's National Cardiothoracic Surgery databases (covering coronary bypass surgery, valve surgery, thoracic surgery, and congenital heart surgery), and the American College of Cardiology's National Cardiovascular Data Registries (including percutaneous coronary intervention, implantable defibrillators, and carotid stenting, as well as diseases such as congenital cardiovascular disease and acute coronary syndromes). State governments have also sponsored clinical registries, such as the New York State Registry for cardiac surgery and percutaneous coronary intervention. Clinical registries have also been established using federal research funding (e.g., the NHLBI-sponsored Dynamic Registry). Industry has sponsored clinical registries on particular drugs or devices (e.g., stent registries), as well as specific diseases (e.g., the NRMI [National Registry of Myocardial Infarction], the ADHERE [Acute Decompensated Heart Failure National] registry, and the REACH [Reduction of Atherothrombosis for Continued Health] registry).

Health system data. Electronic health records and the administrative records of insurers or integrated health plans are valuable sources of observational data for CER. These "found data" are generated in the routine practice of medical care for billing purposes, public reporting, or clinical care and are not produced primarily for research purposes. For instance, the diagnoses and procedures during hospitalization are recorded using the nomenclature of the International Classification of Diseases, Clinical Modification (ICD-9-CM) and reported to public authorities and health insurers. These administrative data are not subject to the standardization and quality control applied to clinical trial or clinical registry data, although these data are usually recorded by trained medical records abstractors and required for provider reimbursement. Some integrated healthcare organizations have additional sources of clinical data, such as drug prescriptions, outpatient claims, and computerized laboratory results. Linkage of several of these data sources can give a very detailed picture of medical care provided to a representative, relatively unselected population of patients.

The advent of fully electronic health records offers the possibility of capturing greatly detailed clinical information about individual patients, such as symptoms, vital signs, and results of imaging studies. There are formidable technical challenges in extracting specific data elements from electronic health records, because clinical notes are typically entered as free text rather than using a controlled vocabulary. Advances in medical informatics, such as studies on natural language processing, will likely facilitate the use of electronic health records for research purposes. Nevertheless, the use of electronic records for research purposes does not overcome the well-recognized fundamental limitations of retrospective chart review studies: namely, that key data may not have been recorded at all; and that the data that were recorded are unlikely to be standardized or quality controlled.

Analysis of observational data. A weakness of all observational CER studies (including analyses of clinical registries, electronic health records, and administrative data) is the absence of randomized assignment of treatments. In contrast to a randomized trial, clinicians and patients represented in observational databases select treatments for a variety of reasons, which may not be recorded in the chart. Treatment selection can lead to differences in patient prognosis between treatment groups, so their subsequent clinical outcomes may differ, even in the absence of a treatment effect. Methods for addressing selection biases, whether due to known or unknown factors, are evolving. One simple step is to restrict the study patient population to newly treated patients and to patients eligible for either treatment; these restrictions narrow any pre-treatment differences between patients receiving alternative therapies (15). Modeling the selection of treatment by using a propensity score or a disease risk score can balance treatment groups on large numbers of measured clinical covariates (16,17). Marginal structural models with inverse weighting by propensity for treatment have been used to estimate the effect on outcomes of treatments that vary over time, such as use of prescription drugs (18). Each of these statistical methods relies on adjusting for clinical characteristics that were recorded in the data and, therefore, may not adjust fully for clinical factors that were not recorded (e.g., patient frailty or socioeconomic status) or that are difficult to capture (e.g., degree of social support). Instrumental variable methods and multilevel analyses have been used in an attempt to adjust for unmeasured confounders (19–21). These approaches identify variables that are strongly related to the likelihood of receiving a specific treatment, but do not directly affect clinical outcomes (e.g., inclusion of specific drugs in a formulary, distance to a referral hospital). The workshop participants recognized the need for further advances in statistical methods in order to conduct high-quality CER that minimizes confounding and residual selection bias in observational treatment comparisons.

Systematic reviews and decision models. Systematic reviews and meta-analyses of existing effectiveness and safety data are an important tool for CER. Systematic reviews can identify evidence gaps, including a lack of evidence or unclear evidence for an important clinical question, which suggest a new trial is warranted (22) and when there is "too much evidence," which suggests further trials may be unnecessary (23). Network meta-analysis (mixed treatment comparison meta-analysis) may be used to compare interventions even when direct "head-to-head" trials are not available (24). This new statistical approach is promising, but it is still being examined for its validity as a method to compare treatments.

The potential for treatment efficacy to vary significantly according to patient characteristics (e.g., age, sex, diabetes, or genetic markers) is an important dimension of CER, as

it provides evidence to "personalize" treatment decisions. Very large sample sizes are required to investigate potential variation in treatment effects across populations, which can be accomplished by pooling individual patient-level data from several trials. Further research could address the practical and methodologic challenges of collaborative studies that pool data from multiple studies of cardiovascular treatments.

Decision models and simulation studies are also valuable methods to foster CER. These techniques start with the best available evidence on a clinical question and typically rely on the results of systematic reviews, meta-analyses, and pooling studies to provide the needed data. Models and simulation studies are particularly well suited to identify evidence gaps and the value of information by pinpointing which clinical factors have the greatest impact on the clinical effectiveness and cost-effectiveness of treatment alternatives. Studies can subsequently be designed to address the most important evidence gaps, which will be particularly valuable in meeting the CER objective of "comparing benefits and harms of alternative methods [of care]."

Proposed Framework for CER

The workshop participants proposed a new framework for evidence development (Fig. 1) to foster cardiovascular CER. This framework identifies gaps between the evidence needed for practice and the research available to support it. A "portfolio analysis" of the current state of knowledge, by using clinical guidelines, evidence reviews, and decision models to identify key evidence gaps, could focus CER on key questions through a variety of research designs, including observational studies, clinical registries, and randomized trials. This cycle of portfolio analysis, identification of key evidence gaps, and research addressing these gaps could repeat over time (Table 1).

The proposed framework (Fig. 1) involves a multiple stakeholder process to identify questions for CER and is consistent with recent changes on the political as well as scientific landscape. For example, the creation of the new Patient-Centered Outcomes Research Institute suggests that the model for setting research priorities may be changing from an investigator and industry driven process to one with broader patient and clinician input.

Table 1 General Recommendations to Foster Comparative Effectiveness Research

Conduct portfolio analyses in key clinical areas
Identify knowledge gaps
Strengthen relationships among stakeholders
Leverage the strength of different research methods: trials, registries, simulations, evidence synthesis
Advance analytic methods for comparative effectiveness research
Promote knowledge discovery as part of clinical practice
Foster training and careers in comparative effectiveness research

The proposed framework implies that knowledge gaps could be identified through systematic reviews of existing evidence, meta-analyses, and decision analytic models. These types of studies could be efficient and timely opportunities to focus on important comparative effectiveness studies (Table 1).

The proposed framework also implies that interactions among potential funders and stakeholders, such as the Agency for Healthcare Research and Quality, the Centers for Medicare and Medicaid Services, and the Centers for Disease Control and Prevention could help foster CER focused on cardiovascular disease. The Cochrane Collaboration is an international organization that conducts systematic reviews and promotes methodologic development, and it could assist in assessing evidence gaps and areas of focus for research. The professional societies (e.g., the American College of Cardiology, the American Heart Association, and the Society of Thoracic Surgeons) now operate clinical registries that may be well suited to cardiovascular CER, and they develop clinical guidelines, performance measures, and appropriate-use criteria that could potentially translate CER findings into practice.

Expand the scope of discovery. CER focuses on developing scientific knowledge that will be useful during the course of patient care. The creativity and innovation that characterize investigator-initiated research have served medical science well. However, because the research questions of CER focus on addressing the needs of practitioners and patients, the scope of discovery may need to expand to include not only specific topics of interest to individual investigators and experts, but also topics driven by clinical evidence gaps.

The implications of developing a CER portfolio that focuses on knowledge gaps are far reaching. The evolution of a research question from model to meta-analysis to mega-trial to implementation underscores the contributions of methodologic research as well as research on specific clinical questions. Whereas some clinical questions may be addressed in large randomized clinical trials, other pivotal questions may not be amenable to randomized studies and, therefore, require alternative methods.

Create a culture of research. The broad scope of CER implies similar breadth in thinking about the research enterprise, such that the development of new knowledge about optimal practice becomes an intrinsic part of the healthcare system. Because CER aims to address patient and provider needs, health systems, providers, and patients should embrace the need to perform CER and recognize its value. A culture of research may be fostered when it is recognized that there is uncertainty regarding what constitutes optimal care, so that alternative forms of management may be reasonable and acceptable—the concept of clinical equipoise. While clinical systems and caregivers face challenges incorporating research as part of daily practice, they should recognize that improving the quality of care through discovery and learning from experience should be part of

their professional and institutional obligations. Integrating research into routine care might be further fostered by focusing on quality improvement, aligning incentives within health systems, and, in the larger healthcare enterprise, by establishing policies such as coverage with evidence development and pay for performance.

The collection of observational data as part of a prospective clinical registry is an example of how research can be incorporated into daily practice. Clinical registries cover only specific clinical populations, however, and a broader knowledge base could advance CER. Fully interoperable electronic health records, with promotion of standardized clinical terminology, would facilitate CER, particularly if electronic health records, clinical registries, image repositories, clinical trial data, and longitudinal claims can be linked to create study cohorts. To this end, further development of health informatics and its application to cardiovascular disease is an opportunity to advance CER.

Nurture the national CER workforce. Individuals from many distinct backgrounds are needed to meet the challenges of CER, ranging from skilled investigators to inquisitive clinicians, knowledgeable patients, and practical methodologists. Thus, "team science" is integral to the success of CER. Paralleling the need for more diverse research partners is the opportunity to collaborate with nontraditional research venues and partners, including those usually focused only on care delivery or education, rather than scholarly research.

As CER is relatively new field, there is a great need for more investigators who are well trained in its methods. Training, mentoring, and professional development programs aimed specifically at expanding the pool of investigators skilled in the methods of cardiovascular CER are possible ways to cultivate this field. Furthermore, nontraditional stakeholders may be valuable contributors to the development of the portfolio of CER studies.

Institute of Medicine Priorities

The workshop addressed 2 broad questions. 1) How might the NHLBI foster CER related to cardiovascular disease in general? 2) How might the NHLBI respond to the specific CER priorities in the area of cardiovascular disease identified by the IOM? In this section, we summarize the workshop participants' recommendations on the second broad question—what types of studies could potentially address the IOM priorities (1)? The workshop participants were charged to identify examples of 1 or 2 study ideas for each IOM area.

NHLBI convenes working groups of experts to provide recommendations and input on specific topic areas, which the Institute then carefully reviews. The suggestions of this workshop represent a list of important areas for investigator-initiated and/or Institute-initiated projects. The Institute carefully considers these and other recommendations as it sets its priorities and attempts to maintain a balanced portfolio across its entire mission; no NHLBI funding commitment is made or implied by inclusion of the topics in the report of this workshop. The workshop recognized that the NHLBI is particularly able to organize CER studies free of conflicts of interest related to specific drugs, devices, or management strategies and has great experience in conducting comparative studies of alternative management strategies to treat cardiovascular disease.

The IOM priority areas discussed at the workshop include:

- Compare the effectiveness of treatment strategies for atrial fibrillation including surgery, catheter ablation, and pharmacologic treatment;
- Compare the effectiveness of anticoagulant therapies (e.g., low-intensity warfarin, aspirin, injectable anticoagulants) for patients undergoing procedures;
- Compare the effectiveness of treatment strategies for vascular claudication (e.g., medical optimization, smoking cessation, exercise, catheter-based treatment, open surgical bypass);
- Compare the effectiveness of aggressive medical management and percutaneous coronary interventions in treating stable coronary disease for patients of different ages and with different comorbidities;
- Compare the effectiveness of innovative treatment strategies (e.g., cardiac resynchronization, remote physiologic monitoring, pharmacologic treatment, novel agents such as CRF-2 receptors) for congestive heart failure;
- Compare the effectiveness of different treatment strategies (e.g., modifying target levels for glucose, lipid, or blood pressure) in reducing cardiovascular complications in newly diagnosed adolescents and adults with type 2 diabetes;
- Compare the effectiveness of traditional risk stratification for coronary heart disease and noninvasive imaging (using coronary artery calcium, carotid intima media thickness, and other approaches) on outcomes; and
- Compare the effectiveness of computed tomography (CT) angiography and conventional angiography in assessing coronary stenosis in patients at moderate pre-test risk of coronary artery disease.

Atrial Fibrillation

Atrial fibrillation is a highly prevalent condition associated with increased cardiovascular mortality and a high risk of stroke. The NHLBI has previously supported comparative treatment trials for atrial fibrillation, including the completed AFFIRM (Atrial Fibrillation Follow-Up Investigation of Rhythm Management) trial (4) and the ongoing CABANA (Catheter Ablation Versus Antiarrhythmic Drug Therapy for Atrial Fibrillation) trial (NCT00911508). The NHLBI is also funding a clinical registry of atrial fibrillation in 2 large integrated health plans to address specific comparative

effectiveness research issues, and it will expand on a previous atrial fibrillation registry (25) and epidemiologic studies of atrial fibrillation. Despite these and other studies of atrial fibrillation, key knowledge gaps remain, including: 1) the generalizability of trials to the larger population of patients with atrial fibrillation; 2) whether the treatment approach ought to vary depending on the subtype of atrial fibrillation; and 3) the effect of treatments on expanded outcomes, such as stroke, dementia, heart failure, quality of life, healthcare utilization, and cost-effectiveness.

The workshop participants recommend that 1 approach to fill these gaps could be a robust clinical registry with broader representation of patients with atrial fibrillation, with a particular focus on patients with new onset disease, to better define risks for adverse outcomes in relation to patient characteristics, subtype of atrial fibrillation, biomarkers, and treatment. In conjunction with the registry, a comprehensive decision model of atrial fibrillation management, as outlined in Figure 2, could be used to identify priority areas for additional CER studies, including clinical trials. There is also the potential to build on ongoing studies of other conditions and treatments by adding atrial fibrillation as a secondary outcome measure.

Anticoagulant Therapies

Anticoagulant and antithrombotic therapies are increasingly used to prevent thromboembolism in patients with atrial fibrillation, deep vein thrombosis, or pulmonary embolism. Management of these therapies at the time of invasive procedures and surgery poses a difficult problem in balancing the bleeding risk due to treatment and the thromboembolic risk due to the underlying disease (26). Furthermore, immobilization after surgery promotes venous thromboembolism, which is particularly increased by hip or knee joint replacement procedures. Management decisions have been further complicated by the introduction of several novel anticoagulant drugs that lack specific antidotes and whose anticoagulant intensity cannot be reliably assessed by laboratory tests. The NHLBI is currently supporting the BRIDGE (Effectiveness of Bridging Anticoagulation for Surgery) trial (NCT00786474), and the GIFT (Genetics Informatics Trial of Warfarin to Prevent DVT) trial (NCT01006733), which tests low-intensity warfarin properties following orthopedic surgery.

The workshop participants felt that a large "real-world" registry of patients undergoing specific surgical procedures would be of benefit to document risk factors for bleeding, cardiac events, and thromboembolism. This registry could be used to assess the effect of different treatments on those outcomes, particularly among patients under-represented in randomized trials (Online Fig. 1). Assessment of adverse orthopedic outcomes such as joint hemorrhage, periprosthetic infection, and repeat procedures would provide critical information needed to balance risks and benefits of anticoagulant treatment. The registry could capitalize on substantial practice variations to perform observational treatment comparisons and assess the effects of treatment on cost, quality of life, and cost-effectiveness.

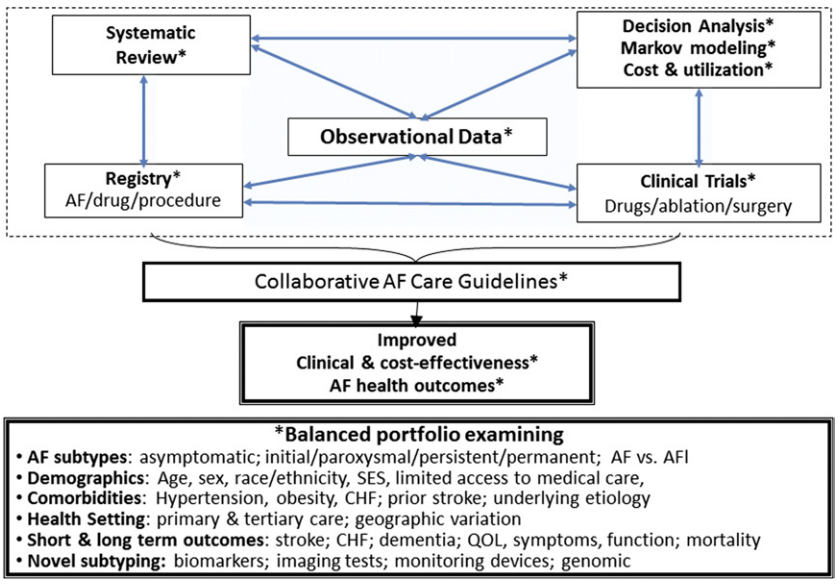


Figure 2 AF CER Evidence and Data Portfolio Management

The general framework for CER (see Fig. 1) is applied to atrial fibrillation (AF). Examples of registries might include an AF clinical care registry, registry of individuals screened but not randomized to a clinical trial, or individuals followed after completion of a clinical trial. An example of an ongoing AF clinical trial is the CABANA (Catheter Ablation Versus Anti-Arrhythmic Drug Therapy for Atrial Fibrillation Trial) (NCT00911508). AFl = atrial flutter; CHF = congestive heart failure; QOL = quality of life; SES = socioeconomic status; other abbreviations as in Figure 1.

JACC Vol. 60, No. 7, 2012
August 14, 2012:569–80

Peripheral Artery Disease

Lower extremity peripheral artery disease is a common, costly condition that is associated with high morbidity and mortality. The pathophysiology of exertional limb claudication is analogous to that of exertional angina pectoris, but claudication has not received as much attention as angina in either clinical investigation or the development of new drugs and devices. There have been relatively few CER studies of claudication treatments, apart from the ongoing NHLBI-sponsored CLEVER (Claudication: Exercise Versus Endoluminal Revascularization) trial (NCT00132743). CLEVER compares exercise therapy with endovascular stenting for treatment of a documented lesion in a specific proximal aortoiliac site that is amenable to stent therapy (27). CLEVER does not, however, address the larger question of whether a strategy of mechanical limb revascularization (using any combination of endovascular techniques and surgery) leads to better clinical outcomes than the alternative strategy of optimal medical management (including exercise, lifestyle modification, and drug therapy). A large practical randomized trial comparing these distinct approaches to claudication could address this priority area (Online Fig. 2). This trial could enroll relatively unselected patients with claudication, whether new onset or after prior revascularization, without requiring prior angiography or any specific anatomic findings. The major outcomes of this trial could be patient functional capacity and quality of life, with secondary outcomes of major clinical complications, cardiovascular risk factor control, medical care utilization and cost, and cost-effectiveness.

Severe limb ischemia, including acute or chronic critical limb ischemia syndromes, represents another major clinical manifestation of peripheral artery disease, in which tissue necrosis is threatened due to reduced resting blood flow. There are newer therapeutic options for severe limb ischemia, but few reliable data on their long-term clinical outcomes or comparative effectiveness. Severe limb ischemia is particularly suitable for the process of priority setting outlined in Figure 1: performing a systematic review of evidence; modeling to identify critical parameters; and collecting observational data on epidemiology, treatment patterns, and determinants of clinically important outcomes. Establishment of a clinical registry of patients with severe limb ischemia could foster CER on this topic by identifying patient, provider, and treatment predictors of outcome. Such a registry could provide the basis for initiating subsequent, targeted clinical trials of evaluation and management strategies for severe limb ischemia.

Stable Ischemic Heart Disease

Coronary artery disease is well recognized as a major health problem in the United States and has been the subject of numerous clinical investigations. The NHLBI has sponsored pivotal CER clinical trials, including the CASS (Coronary Artery Surgery Study) (28) and BARI 2D (Bypass Angioplasty Revascularization Investigation Two, Diabetes) (NCT00006305) (7) to compare coronary revascularization with medical therapy among patients with ischemic heart disease. The NHLBI has also sponsored trials comparing bypass surgery with coronary angioplasty (BARI [29] and EAST [Emory Angioplasty Versus Surgery Trial] [9]). Despite extensive investigation in this field, numerous knowledge gaps persist about optimal management of patients with stable ischemic heart disease. In particular, prior clinical trials have required knowledge of the coronary anatomy prior to randomization, but the decision to perform an invasive coronary angiogram has often been tantamount to the decision to perform coronary revascularization. The NHLBI just announced funding for the ISCHEMIA (International Study of Comparative Health Effectiveness with Medical and Invasive Approaches), a trial that will randomize patients with stable coronary disease and objective evidence of myocardial ischemia to an invasive strategy of early coronary angiography or a conservative strategy of initial optimal medical therapy, with angiography reserved for development of refractory symptoms or a clinical event (Online Fig. 3), in order to address a need for further evidence on this important decision point.

In addition to a large practical trial of coronary revascularization, many questions remain about how to define and deliver optimal medical management for patients with stable ischemic heart disease. In particular, it has been challenging to promote drug adherence and behavior change (diet, exercise, smoking cessation) in the setting of a busy outpatient practice. One approach to consider is to apply insights from behavioral economics to investigate the effect of economic incentives to clinicians (e.g., structuring of payment) or patients (e.g., copayments for drugs or visits, the costs of improving exercise and diet) on clinical outcomes.

As a third consideration, formal analysis of evidence gaps and opportunities in stable ischemic heart disease using the processes of evidence review, model building, and analysis of clinical registries (Fig. 1) would allow the identification of additional opportunities for CER in this area.

Heart Failure

Heart failure continues to be the most common reason for hospital admission among Americans 65 years of age and older, and the prevalence of heart failure has continued to rise, even though other forms of heart disease have been declining. The NHLBI has sponsored numerous CER investigations in heart failure, including the SOLVD (Studies of Left Ventricular Dysfunction) trial (30), the SCD-HeFT (Sudden Cardiac Death in Heart Failure) (NCT00000609) trial (11), and, more recently, the STICH (Surgical Treatment of Ischemic Heart Failure) (NCT00023595) trial (12). Many other studies of the effects of drugs and devices on clinical

outcome in patients with heart failure have been sponsored by industry.

A hallmark of heart failure is that many patients are frequently rehospitalized to treat exacerbations of the disease, at great expense to the system and considerable distress to patients and their families. There are major gaps in knowledge about how to address this problem, so the workshop participants proposed that 1 study to consider is enhanced disease management and transition of care with monitoring (remote or biomarker) as a means of improving clinical outcomes. A large practical trial (Fig. 3) could enroll unselected patients with heart failure at the time of hospital discharge and randomize them either to usual care or to enhanced disease management with tailored therapy guided by remote monitoring (e.g., weight, heart rate, blood pressure, biomarkers). An associated registry of patients with heart failure could be established in conjunction with the randomized trial to collect additional data on the full spectrum of patients with heart failure and thereby assess the generalizability of the trial results.

The proposed study could use a cluster randomized design (31), in which clinical sites rather than individuals are randomized. The cluster randomized design is well suited to evaluation of interventions that target behavior or processes of care. Cluster randomized trials are particularly useful when it is difficult to conceal the nature of the intervention from the clinic staff and to mask patients to the intervention. The novel research designs that may be required for CER entail unique challenges (32). Despite such challenges, the workshop participants encouraged use of such novel research designs, as they provide several advantages, including simpler patient recruitment and "real-life" data directly applicable to clinical practice.

Diabetes

The incidence of diabetes continues to rise, driven in large part by the epidemic of obesity in the United States. Most patients with diabetes die of heart disease, yet the optimal approach to prevention and treatment of cardiovascular disease in patients with diabetes is not well established. Despite the recent publications of the results of the ACCORD (Action to Control Cardiovascular Risk in Diabetes Trial) (33), ADVANCE (Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation) (34), VADT (Veterans Affairs Diabetes) trial (35), long-term follow-up from UKPDS (United Kingdom Prospective Diabetes Study) and Steno 2 Study (36), and the ongoing Look AHEAD (Action for Health in Diabetes) trial (NCT00017953), the comparative effectiveness of tight glycemic control versus more liberal control of diabetes, of lifestyle management versus early drug treatment, of different initial drug treatments, and of different systems of care are all uncertain in patients with *newly* diagnosed diabetes. The workshop participants felt that the IOM priority could potentially be addressed by the following 2 trials.

The first trial could compare 2 strategies for initial treatment in patients with newly diagnosed type 2 diabetes: intensive lifestyle intervention versus immediate metformin (Online Fig. 4). A third arm of this trial could include both interventions. A second trial could focus on the comparative

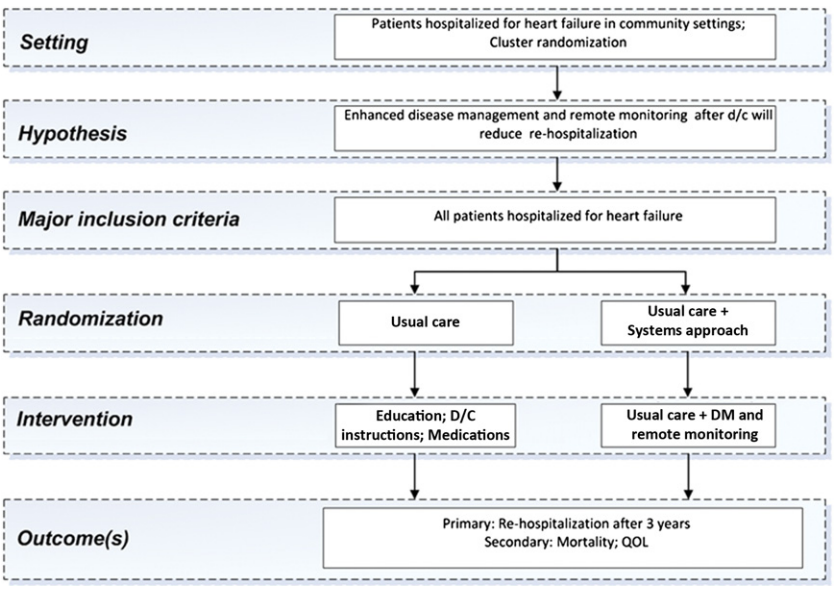


Figure 3 Proposed Cluster Randomized Trial for the Management of Heart Failure

DC = discharge; DM = disease management; QOL = quality of life.

effectiveness on cardiovascular outcomes of using different classes of glucose-lowering drugs as second agents in patients in whom metformin monotherapy fails (Online Fig. 5). Metformin therapy is the current evidence-based standard for initial pharmacotherapy, but multiple medications are usually needed to maintain durable hemoglobin A1C control. In this trial, participants would be randomized to second stage treatment with incretin mimetics versus a thiazolidinedione versus a secretagogue, with insulin withheld for third-line use.

Potential outcomes in each study could include noninvasive measures of atherosclerosis (e.g., coronary artery calcium), progression of diabetes, quality of life, and satisfaction with treatment. In light of unexpectedly higher rate of total mortality in the more intensively treated group in ACCORD, any early differences in subclinical disease markers or surrogate outcomes ought to be confirmed by subsequent assessments of clinical outcomes. Clinical cardiovascular events (cardiac death, myocardial infarction, and stroke) could be potential outcomes as well, although the group recognized that potentially low event rates might require large sample sizes and long follow-up. Depending on the outcome and participant characteristics, this simple trial could range from as few as 1,000 to 2,000 to as many as 40,000 to 50,000 participants.

A registry could also be useful for post-marketing surveillance of cardiovascular disease events among patients prescribed drugs to treat diabetes.

Risk Stratification for Coronary Artery Disease

The current paradigm for prevention of coronary disease in asymptomatic adults is based on individual risk assessment using a standard risk predictor (such as the Framingham Risk Score) and intervening with drug therapy among the individuals identified as being at highest cardiovascular risk (37). The optimal management for the patient at intermediate risk in this paradigm remains uncertain. Additional risk stratification of individuals at intermediate cardiac risk, with drug treatment of individuals with "positive markers," may substantially improve outcomes. There are many candidate risk markers that could be applied at the population level, but coronary calcification on CT is quite promising, as it provides high levels of incremental information and risk reclassification. Consequently, the workshop participants considered a large trial of coronary calcium screening among individuals found to be at intermediate risk on conventional evaluation (Online Fig. 6), with individuals found to have elevated coronary calcium receiving intensive drug therapy and lifestyle modification, and individuals without elevated coronary calcium scores undergoing lifestyle modifications alone. The trial could follow patients for up to 5 years, with the primary outcome of major cardiac events. Secondary outcomes could include quality of life, adherence to drug and lifestyle management, cost, and cost-effectiveness.

The current paradigm of individual risk assessment includes pharmacological therapy for those with elevated Framingham risk scores, but most of the high-risk group will not have a cardiovascular event in the subsequent 10 years. Additional testing of individuals with high Framingham risk scores might identify a subgroup that would not benefit sufficiently to justify lifelong pharmacological therapy. Optimal management of these individuals might be improved by imaging to document the severity of the underlying disease process. The workshop participants also suggest that a comprehensive evidence review and decision modeling of the application of imaging the disease substrate (e.g., by CT coronary angiography or coronary calcium measurement) could help to further stratify risk and guide therapy. This investigation would fit within the framework outlined in Figure 1 and assist in determining the potential value of larger CER studies to address this question.

CT Coronary Angiography

Coronary angiography can now be performed noninvasively with recent generation CT scanners, and in several case series, CT angiography has demonstrated high sensitivity and good specificity when compared with invasive coronary angiography as a reference standard (38,39). These data suggest that CT angiography may be very useful in the evaluation of patients with symptoms of coronary disease. The effect of CT angiography on clinical outcomes is uncertain, however, because visualization of coronary obstructions may well lead to unnecessary or inappropriate coronary revascularization, the value of detecting incidental noncardiac findings is unknown, and the ionizing radiation from CT scanning may lead to adverse events. Consequently, outcome-based studies of the comparative effectiveness of CT coronary angiography and alternative diagnostic strategies would address an important gap in the evidence. The NHLBI is currently funding the PROMISE (Prospective Multicenter Imaging Study for the Evaluation of Chest Pain) trial (NCT01174550), which is randomizing symptomatic patients suspected of having coronary artery disease to either usual stress testing (functional) or CT angiography (anatomic).

Another study to address this IOM priority could be a large clinical trial of patients with symptoms suggestive of coronary disease without high risk features, in which patients would be randomized to either invasive coronary angiography or to CT coronary angiography after a stress test that had either inconclusive or "not high risk" results (Online Fig. 7). Patients in the invasive angiography arm of the study would receive coronary revascularization according to current usual care, whereas in the CT angiography arm of the study, only patients with specific anatomic findings (left main disease, severe 3-vessel disease) would be recommended to receive coronary revascularization (and invasive angiography if needed to further define coronary anatomy). The primary endpoint of this trial study would be major

cardiovascular events (cardiac death, myocardial infarction, stroke), for which the noninferiority of CT angiography would be tested. Secondary endpoints could include quality of life, cost, and cost-effectiveness, for which superiority of CT angiography would be tested. The workshop participants suggest that any research infrastructure created for the proposed trial could be used as an "advanced cardiovascular imaging network" to conduct efficiently other CER studies of imaging, as has been undertaken by the Canadian Atherosclerosis Imaging Network and the Medical Imaging Trials Network of Canada.

The workshop also identified a clinical registry of CT coronary angiography procedures, ideally as an extension of the ongoing National Cardiovascular Disease Registries sponsored by the American College of Cardiology, as a further opportunity to promote CER on this topic.

Conclusions

The recent recognition of the importance of comparative effectiveness research places increasing emphasis on studies that directly inform and improve patient care. The Workshop on Cardiovascular Comparative Effectiveness Research was designed to propose approaches for the NHLBI to consider, using both an overall framework for CER and specific study designs as examples.

The approach to actionable research outlined by workshop participants implies a cycle of research and its application (Fig. 1). The 4 important linked steps in this cycle are: 1) the prospective articulation of research questions based on identifying gaps in knowledge about optimal patient care, incorporating input of stakeholders, including patients; 2) the development of evidence by a variety of research methods to address key evidence gaps; 3) application of the evidence in practice guidelines and standards of care; and 4) determination whether quality of care and patient outcomes are improved.

Several important points follow logically from this vision, beginning with consideration of a new research paradigm. Following that, relationships can be created to identify important research questions and infrastructure developed to perform studies, such as large-scale registries and CER trials, which can foster translation of research results into practice. The training and development of CER researchers and the creation of a culture of learning healthcare systems are important steps in furthering CER. Finally, clear identification of evidence gaps and key questions will help guide the performance of high-quality research.

The IOM has begun this process by identifying several research priorities in cardiovascular disease. This workshop has advanced this process by proposing research ideas to address each of the 8 IOM priority areas. These range from systematic reviews, to secondary analyses of existing data, to registries, to large-scale clinical trials; each could add important evidence needed to improve patient care. The workshop participants hope that these recommendations will provide valuable information to investigators and funding agencies as they seek to advance the nation's commitment to cardiovascular comparative effectiveness research.

Acknowledgments

The authors would like to acknowledge the NHLBI staff, Drs. Denise Bonds and Jean Olson, as well as student volunteer, Mr. Jared Lucas, for their efforts in planning and conducting of this workshop.

Reprint requests and correspondence: Dr. Mark A. Hlatky, Stanford University School of Medicine, HRP Redwood Building, Room T150, 259 Campus Drive, Stanford, California 94305-5405. E-mail: hlatky@stanford.edu.

REFERENCES

1. IOM (Institute of Medicine). Initial National Priorities for Comparative Effectiveness Research. Washington, DC: The National Academies Press, 2009.
2. Concato J, Peduzzi P, Huang G, O'Leary TJ, Kupersmith J. Comparative effectiveness research: what kind of studies do we need? J Investig Med 2010;58:764–9.
3. The ACCORD Study Group. Long-term effects of intensive glucose lowering on cardiovascular outcomes. N Engl J Med 2011;364:818–28.
4. Wyse DG, Waldo AL, DiMarco JP, et al., for the AFFIRM Investigators. A comparison of rate control and rhythm control in patients with atrial fibrillation. N Engl J Med 2002;347:1825–33.
5. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). JAMA 2002; 288:2981–97.
6. The BARI Investigators. The final 10-year follow-up results from the BARI randomized trial. J Am Coll Cardiol 2007;49:1600–6.
7. The BARI 2D Study Group. A randomized trial of therapies for type 2 diabetes and coronary artery disease. N Engl J Med 2009;360:2503–15.
8. Alderman EL, Bourassa MG, Cohen LS, et al. Ten-year follow-up of survival and myocardial infarction in the randomized Coronary Artery Surgery Study. Circulation 1990;82:1629–46.
9. King SB, Kosinski AS, Guyton RA, Lembo NJ, Weintraub WS. Eight-year mortality in the Emory Angioplasty versus Surgery Trial (EAST). J Am Coll Cardiol 2000;35:1116–21.
10. Hochman JS, Lamas GA, Buller CE, et al., for the Occluded Artery Trial Investigators. Coronary intervention for persistent occlusion after myocardial infarction. N Engl J Med 2006;355:2395–407.
11. Bardy GH, Lee KL, Mark DB, et al., for the SCD-HeFT Investigators. Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure. N Engl J Med 2005;352:225–37.
12. Velazquez EJ, Lee KL, Deja MA, et al., for the STICH Investigators. Coronary-artery bypass surgery in patients with left ventricular dysfunction. N Engl J Med 2011;364:1607–16.
13. Manson JE, Hsia J, Johnson KC, et al., for the Women's Health Initiative Investigators. Estrogen plus progestin and the risk of coronary heart disease. N Engl J Med 2003;349:523–34.
14. Myers WO, Blackstone EH, Davis K, Foster ED, Kaiser GC. CASS Registry long term surgical survival: Coronary Artery Surgery Study. J Am Coll Cardiol 1999;33:488–98.
15. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol 2003;158:915–20.
16. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55.
17. Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic Clin Pharmacol Toxicol 2006;98:253–9.
18. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550–60.

19. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. Annu Rev Public Health 1998;19:17–34.
20. Selby JV, Fireman BH, Lundstrom RJ, et al. Variation among hospitals in coronary angiography practices and outcomes after myocardial infarction in a large health maintenance organization. N Engl J Med 1996;335:1888–96.
21. Johnston SC, Henneman T, McCulloch CE, van der Lann M. Modeling treatment effects on binary outcomes with grouped-treatment variables and individual covariates. Am J Epidemiol 2002; 156:753–60.
22. Dickersin K. To reform U.S. health care, start with systematic reviews. Science 2010;329:516–7.
23. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. JAMA 1992;268:240–8.
24. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005;331:897–900.
25. Go A, Hylek EM, Phillips KA, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk factors In Atrial fibrillation (ATRIA) study. JAMA 2001;285:2370–5.
26. Douketis JD, Berger PB, Dunn AS, et al. The perioperative management of antithrombotic therapy: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). Chest 2008;133:299–339.
27. Murphy TP, Hirsch AT, Ricotta JJ, et al., for the CLEVER Steering Committee. The CLaudication: Exercise Vs. Endoluminal Revascularization (CLEVER) study: rationale and methods. J Vasc Surg 2008;47:1356–63.
28. The CASS Investigators. Coronary Artery Surgery Study (CASS): a randomized trial of coronary artery bypass surgery. Survival data. Circulation 1983;68:939–50.
29. The Bypass Angioplasty Revascularization Investigation (BARI) Investigators. Comparison of coronary bypass surgery with angioplasty in patients with multivessel disease. N Engl J Med 1996;335:217–25.
30. The SOLVD Investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. N Engl J Med 1991;325:293–302.
31. Eldridge S, Ashby D, Bennett C, Wakelin M, Feder G. Internal and external validity of cluster randomised trials: systematic review of recent trials. BMJ 2008;336:876–80.
32. Luce BR, Kramer JM, Goodman SN, et al. Rethinking randomized clinical trials for comparative effectiveness research: the need for transformational change. Ann Intern Med 2009;151:206–9.
33. The Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med 2008;358:2545–59.
34. The ADVANCE Collaborative Group. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. N Engl J Med 2008;358:2560–72.
35. Duckworth W, Abraira C, Moritz T, et al., for the VADT Investigators. Glucose control and vascular complications in veterans with type 2 diabetes. N Engl J Med 2009;360:129–39.
36. Gæde P, Lund-Andersen H, Parving HH, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. N Engl J Med 2008;358:580–91.
37. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2010;56:e50–103.
38. Mowatt G, Cummins E, Waugh N, et al. Systematic review of the clinical effectiveness and cost-effectiveness of 64-slice or higher computed tomography angiography as an alternative to invasive coronary angiography in the investigation of coronary artery disease. Health Technol Assess 2008;12:iii–iv, ix–143.
39. Mark DB, Berman DS, Budoff MJ, et al. ACCF/ACR/AHA/NASCI/SAIP/SCAI/SCCT 2010 expert consensus document on coronary computed tomographic angiography: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents. J Am Coll Cardiol 2010;55:2663–99.

Key Words: clinical trials ■ comparative effectiveness ■ research methods.

APPENDIX

Workshop Participants

Emelia Benjamin, Roger Blumenthal, Robert Califf, Kay Dickersin, Pamela Douglas, Gregg Fonarow, Alan Go, David Goff, Alan Hirsch, Mark Hlatky, Judith Hochman, Udo Hoffman, Elaine Hylek, Julie Miller, Sharon-Lise Normand, Douglas Packer, Eric Peterson, Véronique Roger, J. Sanford Schwartz, Joseph Selby, Harold Sox, Barbara Tilley, Sean Tunis, James Udelson, William Weintraub.

For supplementary figures, please see the online version of this paper.

The Effect of Imaging on the Clinical Management of Breast Pain

Mary Beth Howard, MS[1], Tracy Battaglia, MD MPH[1,2], Marianne Prout, MD MPH[2], and Karen Freund, MD MPH[1,2]

[1]Women's Health Unit, Section of General Internal Medicine, Department of Medicine, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, MA, USA; [2]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.

BACKGROUND: Breast pain is a common complaint to primary care and breast specialists. Literature recommends imaging to provide reassurance of benign etiology. The effect of imaging on reassurance and subsequent healthcare utilization has not been described.
OBJECTIVE: To determine if initial imaging for breast pain reduces subsequent utilization.
DESIGN: Retrospective cohort study at a hospital-based breast health practice.
PATIENTS: Women referred for breast pain from 2006–2009.
MAIN MEASURES: Imaging ordered at initial provider visit; clinical utilization, defined as the number of follow-up visits, diagnostic imaging studies, and biopsies completed within 12 months following initial visit.
KEY RESULTS: Sixty-percent of women were age 40 or younger, 87% were from racial/ethnic minority groups. Twenty-five percent had imaging ordered at initial visit. Of those who received initial imaging, 75% had normal radiographic findings, yet 98% returned for additional evaluation. In adjusted analyses, women with initial imaging had increased clinical services utilization (OR 25.4, 95% CI: 16.7, 38.6). Women with normal clinical breast exams who received initial imaging exhibited increased odds for subsequent clinical services utilization (OR 23.8, 95% CI: 12.9, 44.0). Six cancers were diagnosed; imaging in the absence of clinical breast exam abnormalities did not result in any cancer identification.
CONCLUSIONS: Initial imaging for women with breast pain increased the odds of subsequent clinical utilization and did not increase reassurance in ruling out malignancy.

KEY WORDS: breast pain; mammography; breast cancer.
J Gen Intern Med 27(7):817–24
DOI: 10.1007/s11606-011-1982-4

BACKGROUND

Women often seek medical attention for breast pain due to concerns of breast cancer.[1–3] Breast pain accounts for 45–70% of breast-related complaints in the primary care setting.[4–6] Given that breast pain as a sole complaint has low risk of breast cancer (0-3%),[7–9] reassurance of non-malignancy is appropriate.[2,10,11] Nevertheless, reaching definitive diagnosis in patients with breast pain represents a dilemma, as the causes and treatments of breast pain are inadequately defined.

Current guidelines recommend imaging for breast pain if clinically indicated, such as in conjunction with a palpable mass.[12] Previous research has also recommended imaging in patients in need of reassurance,[9,11,13,14] and suggested that after initial imaging, the majority of women require no intervention following reassurance that evaluation findings are normal.[1,15] No studies have examined the effect that initial imaging for evaluation of breast pain has on provider or patient assurance in ruling out malignancy, as well as subsequent clinical management in women with breast pain.

The aim of the present study was to determine how imaging impacts clinical management of breast pain. We assessed whether initial imaging increases reassurance in ruling out breast cancer, as measured by subsequent clinical utilization. We posited that reassurance in ruling out breast cancer would be reflected in reduced subsequent clinical utilization.

Received August 15, 2011
Revised December 2, 2011
Accepted December 14, 2011
Published online January 31, 2012

METHODS

Study Setting and Population

We conducted a retrospective chart review of women referred for breast pain to internists practicing in a hospital-based diagnostic breast health practice at an academic medical center from January 1, 2006 to December 31, 2009. This specialty practice includes internists trained in breast health and a triage protocol that results in the majority of benign referrals triaged to internists.[16] Referrals are scheduled with a provider who follows a woman longitudinally through subsequent breast care received. Data were abstracted from the electronic medical record (EMR) or scheduling system. This study was approved by the Institutional Review Board of Boston University School of Medicine.

Measures

Initial imaging was defined as the completion of a physician-ordered diagnostic mammogram, ultrasound, or magnetic resonance imaging (MRI) within 3 months of initial clinical visit. Screening mammograms were distinguished from diagnostic mammograms through provider-input orders in the EMR. Screening mammograms were excluded from this and our outcome metric.

The outcome of interest was clinical services utilization. Clinical services utilization served as a proxy for reassurance, and was quantified as the number of subsequent diagnostic services completed in the 12 months following initial provider visit. The following discrete clinical variables were included in defining clinical utilization: (1) additional imaging tests completed (diagnostic mammograms, ultrasounds, or MRIs) (2) biopsies completed (fine needle aspiration, core needle, and excisional biopsies), and (3) additional clinical visits to a breast specialist over 12 months of follow-up. These three variables were summed to yield a summary "clinical services utilization score." This summary measure was categorized into 3 levels: 0: no further clinical utilization, 1: one additional follow-up measure, and 2: two or more additional follow-up measures, all over 12 months of follow-up. The components of clinical utilization were also analyzed individually as dichotomous variables: any receipt of additional diagnostic imaging, any biopsy completed, and categorical number of additional visits (no additional visits, one additional visit, and two or more additional visits), all over 12 months of follow-up.

The number of cancers diagnosed during 12 months of follow-up was descriptively analyzed, including clinical presentation, clinical evaluation and diagnostic testing, and timing of diagnostic testing and diagnoses. Cancer diagnoses were confirmed from pathology reports in the EMR. All women in the study were cross-referenced with the Boston Medical Center Cancer Registry to ensure no cancer diagnoses were missed.

Since abnormal clinical breast exam results could confound analyses by clinical indication, the study population was stratified based on the following categories of clinical exams: (1) normal clinical breast exam, (2) mass on clinical breast exam and (3) abnormality other than mass on clinical breast exam (including breast skin changes, nipple changes, nipple discharge). Stratification in this way allowed for restricted analyses on women with normal clinical breast exams.

Covariates in analyses included demographics and risk factors for breast cancer: age ($\geq$ 40 or <40 years), language (English or non-English speaking), race/ethnicity (White, Black, Hispanic, or Other), insurance (private, public, or no insurance), family history of breast cancer, current hormone therapy use, and current oral contraceptive use. Race/ethnicity minority status was included as it has been associated with delays in cancer screening,[17] diagnosis,[18,19] and treatment.[19–21] Year of referral was included in analyses to account for potential secular trends in imaging and diagnostic utilization.

Statistical Analyses

Demographic differences between women with and without initial imaging were identified using the chi-square test or t-test. Subsequent clinical services utilization was compared between the women who received initial imaging and those who did not. These associations were examined within each of the three strata of clinical breast exam results. Unadjusted logistic regressions determined the odds of each measure of clinical utilization in women who received initial imaging compared to those who did not receive initial imaging.

Multivariate ordinal logistic regression models assessed the effect of initial imaging on subsequent clinical utilization, controlling for demographic and clinical variables. Multivariate models were applied to the study population as a whole and the three clinical breast exam result strata. Variables that were not associated with the outcome at the $p<0.05$ level and variables that did not change effect estimates by greater than 10% were removed from the model.

Sensitivity analyses were conducted to rule out alternative explanations of findings. For women with normal clinical breast exams, we compared the subgroup with no initial imaging to the subgroup with normal initial imaging results (Breast Imaging-Reporting and Data System (BIRADS) 1 or 2). Using age as a proxy for menopausal status, we conducted a stratified analysis of women less than and equal or greater than 50 years of age. Provider seen at initial visit was included in analyses to account for provider-specific practices in managing breast pain. Because of the small number of patients seen by some providers, cluster analysis was not possible. Instead, we stratified the analyses by one outlier provider with a higher rate of initial imaging compared to the other providers. All data were analyzed using Statistical Analysis System version 9.1 (SAS Institute, Cary, NC).

RESULTS

Breast pain accounted for 32% of new patient referrals seen by internal medicine breast providers from January 2009–December 2009. The mean age was 39 ±13 years with 60% of women under age 40, 87% were of minority race/ethnicity, 55% were English speaking, and 73% had no insurance or public health insurance (Medicaid or Medicare) (Table 1). Twenty percent of women reported current oral contraceptive use, and 2% reported current postmenopausal

Table 1. Association of Demographic and Clinical Variables Associated with Receipt of Initial Imaging for Breast Pain

	Total	Imaging within 3 months of referral		
	N (%)	Yes, *N* (%)	No *N* (%)	
	N=916	*N*=229 (25.0)	*N*=687 (75.0)	*P*-values
Age				
<40 years	548 (59.8)	142 (62.0)	406 (59.1)	
≥40 years	368 (40.2)	87 (38.0)	281 (40.9)	$p=0.44$
Race/Ethnicity				
White	121 (13.2)	43 (18.8)	78 (11.4)	
Black	298 (32.5)	76 (33.2)	222 (32.3)	
Hispanic	376 (41.1)	85 (37.1)	291 (42)	
Other	121 (13.2)	25 (10.9)	96 (14)	$p=0.023$
Language				
English speaking	507 (55.4)	144 (62.9)	363 (53)	
Non-English speaking	409 (45.7)	85 (37.1)	324 (47)	$p=0.0081$
Insurance				
None	177 (19.3)	45 (19.7)	132 (20)	
Public	492 (53.7)	125 (54.6)	367 (53)	
Private	247 (27.0)	59 (25.8)	188 (27)	$p=0.89$
Clinical Breast Exam Results				
Mass	111 (12.1)	75 (32.8)	36 (5)	
Abnormality other than mass	293 (32.0)	79 (34.5)	214 (31)	
Normal	512 (55.9)	75 (32.8)	437 (64)	$p<0.0001$
Family History of Cancer				
Yes	195 (21.3)	61 (26.2)	134 (20)	
No	721 (78.7)	168 (73.4)	553 (80)	$p=0.02$
Oral Contraceptives				
Yes	197 (21.5)	59 (25.8)	138 (20)	
No	719 (78.5)	170 (74.2)	549 (80)	$p=0.07$
Hormone Therapy				
Yes	16 (1.7)	7 (3.1)	9 (1)	
No	900 (98.3)	222 (96.9)	678 (99)	$p=0.11$
Year of referral				
2006	225 (24.6)	40 (17.5)	185 (26.9)	
2007	215 (23.5)	53 (23.1)	162 (23.6)	
2008	198 (21.6)	52 (22.7)	146 (21.2)	
2009	278 (30.3)	84 (36.7)	194 (28.2)	$p=0.31$
Provider				
1	156 (17.0)	27 (11.8)	129 (18.9)	
2	187 (20.4)	105 (45.9)	82 (11.9)	
3	228 (24.9)	26 (11.4)	202 (29.4)	
4	12 (1.3)	7 (3.1)	5 (0.7)	
5	19 (2.1)	8 (3.5)	11 (1.6)	
6	21 (2.3)	2 (0.9)	19 (2.8)	
7	60 (6.6)	18 (7.9)	42 (6.1)	
8	233 (25.4)	36 (15.7)	197 (28.7)	$p<0.0001$

hormone therapy use. Over half had a normal clinical breast exam (55%); 12% had a palpable mass noted on clinical breast exam.

Receipt of Initial Imaging

Twenty-five percent of women were referred for diagnostic imaging at initial provider visit (Table 1). Women who received initial imaging were more likely to be white race/ethnicity ($p=0.02$), English-speaking ($p=0.008$), have a mass on clinical breast exam ($p<0.0001$), and have a family history of breast cancer ($p=0.02$) than women who did not receive initial imaging. Receipt of initial imaging was associated with provider seen at initial visit ($p<0.0001$). When the association between provider and initial imaging was stratified by clinical breast exam results, this association was not significant in women with a mass on clinical breast exam ($p=0.06$), but remained in women with normal clinical breast exams ($p<0.0001$).

Subsequent Clinical Services Utilization

Women who received initial imaging were more likely to have subsequent imaging, biopsies, additional visits, and higher clinical services utilization than women who do not receive initial imaging (Table 2). Ninety-eight percent of women who received imaging initially had additional clinical services utilization, versus 26% of women who did not receive imaging ($p<0.0001$). After adjusting for clinical breast exam results, age, family history, and provider, the odds of having a higher level of clinical services utilization for women who received initial imaging were 25.4 (95% CI: 16.7,38.6). While race and language were significantly associated with receipt of initial imaging

Table 2. Association of Initial Imaging for Breast Pain with Subsequent Clinical Utilization

	Imaging within 3 months of referral			Unadjusted OR *(95% CI†)	Adjusted‡ OR(95% CI)
	Yes, *N*=229	No *N*=687	*p*-value*		
	n (%)	*n* (%)			
Follow-up imaging within 12 months					
Yes	126 (55.0)	79 (11.5)	$p<0.0001$	8.8 (6.2,12.5)	7.9 (5.1, 12.2)
No	103 (45.0)	608 (88.5)			
Follow-up biopsy within 12 months					
Yes	38 (16.6)	27 (3.9)	$p<0.0001$	6.0 (3.4, 10.7)	2.1 (1.1, 4.1)
No	191 (83.4)	660 (96.1)			
Sum of visits over 12 months					
0	90 (39.3)	514 (74.8)	$p<0.0001$	4.1(3.1, 5.6)	2.5 (1.8, 3.6)
1	93 (40.6)	119 (17.3)			
2+	46 (20.0)	54 (7.9)			
Clinical services utilization score§					
0	6 (2.6)	509 (74.1)	$p<0.0001$	35.3 (24.1, 51.6)	25.4 (16.7, 38.6)
1	62 (27.1)	121 (17.6)			
2+	161 (70.3)	57 (8.3)			

** OR Odds ratio*
† CI Confidence interval
‡ Adjusted for clinical breast exam results (normal, mass, or other abnormality), age, family history, and provider.
§ Clinical services utilization score: sum of diagnostic follow-up imaging, biopsies, and visits over 12 months of follow-up

(Table 1), these variables did not change the estimate of the effect size in the models and therefore were not included in final models.

When the study population was stratified by clinical breast exam results, the adjusted association between initial imaging and subsequent clinical services utilization remained (Tables 3, 4, and 5). Women with normal clinical breast exams who received initial imaging had 23.8 (95% CI: 12.9, 44.0) times the odds of increased clinical utilization than women who did not receive initial imaging, controlling for age, family history, and provider. Looking at the specific components of utilization (diagnostic imaging, biopsies, and visits), these women had 10.4 (95% CI: 5.5, 19.2) times the odds of receiving additional imaging, 3.7 (95% CI: 1.1, 12.2) times the odds of receiving a biopsy, and 2.3 (95% CI: 1.4, 3.9) times the odds of having additional visits.

Eight providers delivered care to women in the study sample. Four providers saw the majority (87%) of women

Table 3. Association of Initial Imaging for Breast Pain with Subsequent Clinical Utilization in Women with Mass on Clinical Breast Exam, *n*=111

	Imaging within 3 months of referral		Unadjusted OR* (95% CI†)	Adjusted‡ OR (95% CI)
	Yes, *n*=75	No, *n*=36		
	n (%)	*n* (%)		
Follow-up imaging within 12 months				
Yes	45 (60.0)	11 (30.5)	3.1 (1.3, 7.10)	8.5 (2.5, 28.8)
No	30 (40.0)	25 (69.5)		
Follow-up biopsy within 12 months				
Yes	26 (34.7)	13 (36.1)	3.7 (2.3, 6.0)	1.2 (0.5, 3.3)
No	49 (65.3)	23 (63.9)		
Sum of visits over 12 months				
0	21 (28.0)	25 (69.4)	4.4 (2.0, 10.0)	6.5 (2.5, 16.7)
1	34 (45.3)	5 (13.9)		
2+	20 (26.7)	6 (16.7)		
Clinical services utilization score§				
0	2 (2.7)	22 (61.1)	31.7 (11.5, 87.3)	37.6 (12.2, 116.0)
1	11 (14.7)	8 (22.2)		
2+	62 (82.7)	6 (16.7)		

** OR Odds ratio*
† CI Confidence interval
‡ Adjusted for age, family history, and provider
§ Clinical services utilization score: sum of diagnostic follow-up imaging, biopsies, and visits over 12 months of follow-up

Table 4. Association of Initial Imaging for Breast Pain with Subsequent Clinical Utilization in Women with Abnormalities Other than Mass on Clinical Breast Exam, *n*=293

	Imaging within 3 months of referral		Unadjusted OR* (95% CI†)	Adjusted‡ OR (95% CI)
	Yes, *n*=79	No, *n*=214		
	n (%)	*n* (%)		
Follow-up imaging within 12 months				
Yes	41 (51.9)	23 (10.7)	8.5 (4.6, 15.8)	7.6 (3.3, 17.1)
No	38 (48.1)	191 (89.3)		
Follow-up biopsy within 12 months				
Yes	7 (8.9)	5 (2.3)	4.1 (1.3, 13.2)	1.8 (0.4, 7.6)
No	72 (91.1)	209 (97.6)		
Sum of visits over 12 months				
0	39 (49.4)	163 (76.2)	3.3 (1.9, 5.6)	2.1 (1.1, 4.2)
1	27 (34.2)	39 (18.2)		
2+	13 (16.5)	12 (5.6)		
Clinical services utilization score§				
0	1 (1.3)	162 (75.7)	39.3 (19.8, 77.9)	28.4 (13.2, 61.3)
1	30 (38.0)	39 (18.2)		
2+	48 (60.7)	13 (6.1)		

** OR Odds ratio*
† CI Confidence interval
‡ Adjusted for age, family history, and provider
§ Clinical services utilization score: sum of diagnostic follow-up imaging, biopsies, and visits over 12 months of follow-up

during the study. The variability in imaging ordering behaviors varied among both the low- and high-volume providers. The percent of patients receiving initial imaging ranged from 5–31% among providers. Since the subjects seen by Provider 2 made up 46% of the population that received imaging, we performed a sensitivity analysis separating this provider from the other seven providers. The odd ratios of adjusted subsequent imaging was 6.3 (adjusted for age and family history, 95%CI: 2.8, 14.3) compared with the adjusted OR for all other providers (7.7, 95% CI: 4.6, 13.1), suggesting that the findings were not attributable to this provider.

We did not control for imaging results in our analyses, as BIRADS results were significantly associated with all outcome measures. Of the 229 women who received initial imaging, 25% had results that required follow-up (BIRADS

Table 5. Association of Initial Imaging for Breast Pain with Subsequent Clinical Utilization in Women with Normal Clinical Breast Exams, *n*=512

	Imaging within 3 months of referral		Unadjusted OR* (95% CI†)	Adjusted‡ OR (95% CI)
	Yes, *n*=75	No, *n*=437		
	n (%)	*n* (%)		
Follow-up imaging within 12 months				
Yes	40 (53.3)	45 (10.3)	9.4 (5.5, 16.3)	10.4 (5.5, 19.2)
No	35 (46.7)	392 (89.7)		
Follow-up biopsy within 12 months				
Yes	5 (6.7)	9 (2.1)	3.4 (1.1, 10.4)	3.7 (1.1, 12.2)
No	70 (93.3)	428 (97.9)		
Sum of visits over 12 months				
0	30 (40.0)	326 (74.6)	3.7 (2.3, 6.0)	2.3 (1.4, 3.9)
1	32 (42.7)	75 (17.1)		
2+	13 (17.3)	36 (8.2)		
Clinical services utilization score§				
0	2 (2.7)	325 (74.4)	27.4 (15.4, 48.6)	23.8 (12.9, 44.0)
1	23 (30.7)	74 (16.9)		
2+	50 (66.7)	38 (3.7)		

** OR Odds ratio*
† CI Confidence interval
‡ Adjusted for age, family history, and provider
§ Clinical services utilization score: sum of diagnostic follow-up imaging, biopsies, and visits over 12 months of follow-up

Table 6. Initial Clinical Management of Women with Cancer in Cohort of Women with Breast Pain

Case	Age at initial visit	Patient-reported reason for referral	Clinical breast exam results	Initial Imaging	Imaging results	Days from initial visit
1	59	Bilateral breast pain	Mass, left breast	Bilateral diagnostic mammogram	BIRADS* 4	2
2	47	Left breast pain	Mass, left breast	Bilateral diagnostic mammogram and Ultrasound	BIRADS 5	0
3	58	Right breast pain	Mass, right breast	Diagnostic mammogram		0
4	33	Left breast pain	Mass, left breast	Bilateral diagnostic mammogram and Ultrasound	BIRADS 4	77
5	58	Left breast pain	Tenderness, left breast	Unilateral Diagnostic mammogram	BIRADS 1	6
6	59	Bilateral breast pain	Normal	None		0

** BIRADS Breast imaging-reporting and data system*

0, 3, 4, or 5), while 97% went on to receive subsequent diagnostic evaluation (clinical, radiographic, or biopsy). When comparing the 437 women with normal clinical exam and no initial imaging to the 58 women who had a normal exam and initial imaging that revealed normal findings (BIRADS 1 or 2), the adjusted OR remained high at 18.0 (95%CI: 9.4, 59.0), indicating that with normal imaging results, increased subsequent utilization remains. Stratifying women by age greater than 50 years (adjusted OR=12.3, 95% CI: 5.2, 36.8) or less than 50 years (adjusted OR 5.4, 95% CI: 3.3, 8.821) did not show major differences in subsequent additional imaging.

Breast Cancer Diagnoses

Six (0.6%) breast cancers were diagnosed in this study population during the study timeframe, four ductal carcinoma in situ (DCIS) and two invasive ductal carcinoma (IDC, Tables 6 and 7). Cross-referencing with the hospital cancer registry revealed no additional cancer diagnoses in the study cohort. Five of the 6 women initially presented with an abnormal clinical breast exam; 4 with a mass and one with focal tenderness. Four of these women were diagnosed through imaging that was initiated as result of a mass found on clinical breast examination. Table 6 and 7 shows that three of them had timely diagnostic services as a result (Cases 1, 2, 3), while one had almost one year delay due to development of cellulitis and two missed appointments (Case 4). One of these cancers (Case 3) was diagnosed in the breast contralateral to the pain, so that only three of these four cancer diagnoses were concordant with the presenting symptoms. One woman presented with focal tenderness (Case 5) but no discrete mass on initial clinical breast exam and had diagnostic imaging that was read as normal. At a follow-up visit 92 days later, a mass was found on the breast contralateral to the initial site of pain. The patient declined a breast biopsy twice and did not keep one appointment, resulting in a delay in her diagnosis. Only one cancer was diagnosed in a woman who presented with a normal clinical breast examination (Case 6). This case

Table 7. Follow-Up Care of Women with Cancer in Cohort of Women with Breast Pain

Case	Follow-up recommended	Follow-up diagnostic testing	Days from initial visit	Follow-up diagnostic testing results	Cancer site concordant with pain?
1	Right stereotactic biopsy	Right breast stereotactic biopsy	12	DCIS†, right breast	No
2	Referral to breast surgeon	Ultrasound guided core biopsy, left breast	18	IDC‡, left breast, stage 1	Yes
3	Diagnostic Mammography	Unilateral Diagnostic Mammogram, right breast	14	BIRADS 4	
		Right breast stereotactic biopsy	21	IDC, right breast, stage 2	Yes
4	Referral to breast surgeon	Left breast needle localization	223	DCIS, left breast	Yes
5	Follow up with provider in 2 months	Left breast Ultrasound	92	BIRADS 1	
		Bilateral diagnostic mammogram	224	BIRADS 4	
		Bilateral Ultrasound	224	BIRADS 4	
		Right stereotactic needle core biopsy	245	Patient declined biopsy	
		Right breast needle localization	330	DCIS, right breast	No
6	Screening Mammography	Unilateral Mammogram, right breast	16	BIRADS 4	
		Core breast biopsy, right	34	DCIS, right breast	No

** BIRADS Breast imaging-reporting and data system*
† DCIS Ductal carcinoma in situ
‡ IDC Invasive ductal carcinoma

presented with bilateral pain and an age-appropriate screening mammogram (not considered initial imaging in this study) was performed which revealed an incidental finding of Stage 2 IDC. In summary, three of the women had incidental cancer diagnoses (the cancer site did not correspond with the patient-reported area of pain or clinical findings); three had a cancer concordant with their clinical presentation and exam findings.

DISCUSSION

This is the first study to date that measures outcomes in addition to cancer diagnoses in women with breast pain. Our results indicate that initial imaging in the evaluation of breast pain increases subsequent clinical utilization, regardless of clinical breast exam findings. Women who received initial imaging were significantly more likely to undergo additional diagnostic evaluation. Most importantly, women with normal clinical breast exams had increased odds of clinical utilization if they received initial imaging, with no benefit of increased cancer detection. The findings did not change when the data was stratified by age as a proxy for menopausal status, or stratified by one provider with higher rates of initial imaging.

These results support existing data demonstrating a low probability of malignancy in women presenting with breast pain as a primary complaint.[9–11] The number of cancers diagnosed represents 0.6% of the study population, within the range of 0-3% previously documented in the medical literature.[7–9] Three of the 6 cancers were detected with initial imaging showing a lesion that corresponded with a mass at the site of breast pain; one had a negative initial mammogram with imaging three months later finding a contralateral DCIS, while two were detected through screening mammography (not considered diagnostic imaging in analyses). Diagnostic imaging in women with breast pain and normal clinical exams yielded no cancer diagnoses.

Previous studies in women with breast pain have sought to describe the causes, prevalence, and treatment of breast pain. Studies analyzing imaging in evaluation of breast pain have focused on cancer diagnosis as the outcome, and have demonstrated low yield of imaging in the setting of normal findings on clinical examinations.[8–11,14,22] Nevertheless, imaging has been recommended for reassurance purposes,[9] with no data describing its effect on the management of breast pain. By looking at clinical utilization outcomes, we measured the effect imaging has on clinical management of breast pain. Our data show that imaging in the initial evaluation of breast pain leads to increased clinical utilization without increased breast cancer detection. While initial imaging in women with breast pain has been recommended for reassurance purposes, there is significant increased subsequent utilization in women who receive initial imaging, without increased diagnostic yield.

Overutilization of diagnostic imaging is a concern, particularly as healthcare reform demands efforts to curtail overutilization.[23,24] In addition, normal test results do not necessarily lead to reassurance, and in some cases can increase anxiety levels and do harm.[25–27] With efforts to improve health care quality while decreasing costs, it is important to determine if imaging for patients with breast pain is of value in reassuring patients and providers, as reflected in subsequent utilization. The fact that individual provider behaviors vary within the same clinic in the management of women with breast pain and normal clinical breast exams (this variation was absent in women with mass on clinical breast exam), suggests a need for establishment of guidelines for women with breast pain.

Past studies have posited that the goal of imaging in breast pain is to provide reassurance of benign etiology to the patient and provider. This implies that diagnostic certainty of non-malignancy should increase. Previous research has demonstrated a link between diagnostic certainty and provider clinical actions, such that reduced test-ordering behaviors are directly influenced by providers' increased certainty regarding diagnoses.[28–30] Applying this association to our study, with reassurance and diagnostic certainty, subsequent testing should decrease. The increased utilization observed in this study suggests the opposite, that initial imaging does not provide reassurance or increase diagnostic certainty.

Several limitations should be considered when interpreting study findings. Study data did not allow for breast cancer risk adjustment. Using a tool such as the Modified Gail Model[31] was not age-appropriate for all women and there was incomplete data for variables in the tool, including age at first pregnancy, menarche, and menopause. Clinical services provided outside of the institution were not included in analyses. Systematic data to categorize non-malignant diagnoses (i.e. fibroadenoma, cyst) were not available and therefore not included in analyses. Clinical services and additional visits during 12 months of follow-up attributed to complaints other than breast pain could not be ascertained, and therefore we were not able to exclude these visits from analyses.

A potential referral bias exists in this population in that only women with breast pain referred to a specialty practice were included. Providers in this practice have expertise in clinical breast exams and are likely to have a higher sensitivity and specificity of their exams than most primary care providers. Markers of patient concern or anxiety were not collected and therefore could not be controlled for in analyses. Using clinical utilization as a proxy for diagnostic certainty does not elucidate whether the patient or provider is driving increased utilization. Patients who received diagnostic imaging following initial provider visit may have demonstrated a higher level of anxiety or concern than patients who did not. Similarly, providers themselves may

be uncertain about the underlying etiology of breast pain and therefore order additional diagnostic tests. Future studies that prospectively assess anxiety and reasoning for subsequent imaging are needed to address these concerns.

While past studies have indicated the main value of breast imaging in women with painful breasts to be that of reassurance, our results show that initial imaging leads to additional evaluation. Our results support previous research demonstrating that the prevalence of cancer in patients with breast pain is low and suggest that following normal clinical exam, diagnostic imaging is not required to either rule out cancer or provide reassurance in ruling out cancer. As importantly, these results support the critical role of clinical breast exam skills in the evaluation of breast pain.

Contributors: *The authors thank Timothy C. Heeren, PhD, Boston University School of Public Health, for his input on statistical analyses and data presentation. No compensation was received.*

Funding/Support: *This publication was made possible by grant 05-2011-023 from the Avon Foundation Safety Net Funding Initiative Grant and grant KG101421 from the Susan G. Komen Foundation.*

Role of the Sponsors: *The funding organizations were not involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.*

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*

Corresponding Author: *Mary Beth Howard, MS; Women's Health Unit, Section of General Internal Medicine, Department of Medicine, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, 801 Massachusetts Avenue, Suite 470, Boston, MA 02118, USA (e-mail: mbhoward@bu.edu).*

REFERENCES

1. **Millet AV, Dirbas FM.** Clinical Management of Breast Pain: A Review. Obstet Gynecol Surv. 2002;57(7):451–61.
2. **Ganschow P, Norlock F, Jacobs E, Marcus E.** Breast health and vommon nreast problems: a practical approach. Philadelphia: American College of Physicians; 2004.
3. **BeLieu R.** Mastodynia. Obstetrics and gynecology clinics of North America. 1994;21(3):461–77.
4. **Ader D, Browne M.** Prevalence and impact of cyclic mastalgia in a United States clinic-based sample. Am J Obstet Gynecol. 1997;177:126–32.
5. **Leinster SJ, Whitehouse GH, Walsh PV.** Cyclical mastalgia: Clinical and mammographic observations in a screened population. Br J Surg. 1987;74(3):220–2.
6. **Barton MB, Elmore JG, Fletcher SW.** Breast Symptoms among Women Enrolled in a Health Maintenance Organization: Frequency, Evaluation, and Outcome. Ann Intern Med. 1999;130(8):651–7.
7. **Lumachi F, Ermani M, Brandes AA, et al.** Breast complaints and risk of breast cancer. Population-based study of 2,879 self-selected women and long-term follow-up. Biomed Pharmacother. 2002;56(2):88–92.
8. **Smith RL, Pruthi S, Fitzpatrick LA.** Evaluation and Management of Breast Pain. Mayo Clin Proc. 2004;79(3):353–72.
9. **Duijm LEM, Guit GL, Hendriks JHCL, Zaat JOM, Mali WPTM.** Value of breast imaging in women with painful breasts: observational follow up study. Br Med J. 1998;317(7171):1492–5.
10. **Preece P, Baum M, Mansel R, et al.** Importance of mastalgia in operable breast cancer. Br Med J (Clinical Research Edition). 1982;284 (6325):1299–300.
11. **Tumyan L, Hoyt AC, Bassett LW.** Negative predictive value of sonography and mammography in patients with focal breast pain. Breast J. 2005;11(5):333–7.
12. **Morrow M.** The evaluation of common breast problems. Am Fam Physician. 2000;61(8):2371–8.
13. **Ader DN, Shriver CD.** Cyclical mastalgia: prevalence and impact in an outpatient breast clinic sample. J Am Coll Surg. 1997;185(5):466–70.
14. **Brand IR, Sapherson DA, Brown TS.** Breast imaging in women under 35 with symptomatic breast disease. The British Journal of Radiology. 1993;66(785):394–7.
15. **Klimberg V.** Etiology and management of breast pain. In: **Harris J, Lippman M, Morrow M, Hellman S, eds.** Diseases of the Breast, vol. 99-106. Philadelphia: Lippincott-Raven; 1996.
16. **Stone MD, Norton S, Mendez JE, Hirsch E.** Positive impact of a breast-health triaging system on breast-care access and physician satisfaction. Am J Surg. 2007;194(4):482–7.
17. **Swan J, Breen N, Coates RJ, Rimer BK, Lee NC.** Progress in cancer screening practices in the United States. Cancer. 2003;97(6):1528–40.
18. **Chang S, Kerlikowske K, Napoles-Springer A, Posner S, Sickles E, PerezStable E.** Racial differences in timeliness of follow-up after abnormal screening mammography. Cancer. 1996;78(7):1395–402.
19. **Gwyn K, Bondy ML, Cohen DS, Lund MJ, Liff JM, Flagg EW, et al.** Racial differences in diagnosis, treatment, and clinical delays in a population-based study of patients with newly diagnosed breast carcinoma. Cancer. 2004;100(8):1595–604.
20. **Gross CP, Smith BD, Wolf E, Andersen M.** Racial disparities in cancer therapy. Cancer. 2008;112(4):900–8.
21. **Shavers VL, Brown ML.** Racial and ethnic disparities in the receipt of cancer treatment. J Natl Canc Inst J Natl Canc Inst. 2002;94(5):334–57.
22. **Kerin M, O'Hanlon D, Khalid A, Kent P, McCarthy P, Given H.** Mammographic assessment of the symptomatic nonsuspicious breast. Am J Surg. 1997;173(3):181–4.
23. **Hendee WR, Becker GJ, Borgstede JP, et al.** Addressing overutilization in medical imaging. Radiology. 2010;257(1):240–5.
24. **Levin DC, Rao VM.** Turf wars in radiology: Other causes of overutilization and what can be done about it. J Am Coll Radiol. 2004;1(5):317–21.
25. **McDonald IG, Daly J, Jelinek VM, Panetta F, Gutman JM.** Opening Pandora's box: the unpredictability of reassurance by a normal test result. Br Med J. 1996;313(7053):329–32.
26. **Fitzpatrick R.** Telling patients there is nothing wrong. Br Med J. 1996;313(7053):311–2.
27. Screening for breast cancer: U.S. Preventive Services Task Force Recommendations. Ann Intern Med. 2009;151(10):I–44.
28. **Lutfey KE, Link CL, Marceau LD, et al.** Diagnostic certainty as a source of medical practice variation in coronary heart disease: results from a cross-national experiment of clinical decision making. Med Decis Making. 2009;29(5):606–18.
29. **Balsa AI, Seiler N, McGuire TG, Bloche MG.** Clinical uncertainty and healthcare disparities. *Am J Law Med.* 2003;29(207).
30. **Lutfey KE, Link CL, Grant RW, Marceau LD, McKinlay JB.** Is certainty more important than diagnosis for understanding race and gender disparities?: An experiment using coronary heart disease and depression case vignettes. Health policy (Amsterdam, Netherlands). 2009;89(3):279–87.
31. **Gail MH, Brinton LA, Byar DP, et al.** Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Canc Inst J Natl Canc Inst. 1989;81 (24):1879–86.

Thrombosis Research 131 (2013) 31–36

Contents lists available at SciVerse ScienceDirect

Thrombosis Research

journal homepage: www.elsevier.com/locate/thromres

Regular Article

Identifying Major Hemorrhage with Automated Data: Results of the Veterans Affairs Study to Improve Anticoagulation (VARIA)

Guneet K. Jasuja [a,b,*], Joel I. Reisman [a], Donald R. Miller [a,b], Dan R. Berlowitz [a,b,c], Elaine M. Hylek [a,c], Arlene S. Ash [a,d], Al Ozonoff [a,e], Shibei Zhao [a], Adam J. Rose [a,c]

[a] Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, Bedford, MA
[b] Department of Health Policy and Management, Boston University School of Public Health, Boston, MA
[c] Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA
[d] Department of Quantitative Health Sciences, Division of Biostatistics and Health Services Research, University of Massachusetts School of Medicine, Worcester, MA
[e] Clinical Research Center, Boston Children's Hospital, Boston, MA

ARTICLE INFO

Article history:
Received 12 July 2012
Received in revised form 29 September 2012
Accepted 22 October 2012
Available online 13 November 2012

Keywords:
anticoagulants
hemorrhage
outcome assessment
methods

ABSTRACT

Introduction: Identifying major bleeding is fundamental to assessing the outcomes of anticoagulation therapy. This drives the need for a credible implementation in automated data for the International Society of Thrombosis and Haemostasis (ISTH) definition of major bleeding.
Materials and Methods: We studied 102,395 patients who received 158,511 person-years of warfarin treatment from the Veterans Health Administration (VA) between 10/1/06-9/30/08. We constructed a list of ICD-9-CM codes of "candidate" bleeding events. Each candidate event was identified as a major hemorrhage if it fulfilled one of four criteria: 1) associated with death within 30 days; 2) bleeding in a critical anatomic site; 3) associated with a transfusion; or 4) was coded as the event that precipitated or was responsible for the majority of an inpatient hospitalization.
Results: This definition classified 11,240 (15.8%) of 71, 338 candidate events as major hemorrhage. Typically, events more likely to be severe were retained at higher rates than those less likely to be severe. For example, Diverticula of Colon with Hemorrhage (562.12) and Hematuria (599.7) were retained 46% and 4% of the time, respectively. Major, intracranial, and fatal hemorrhage were identified at rates comparable to those found in randomized clinical trials however, higher than those reported in observational studies: 4.73, 1.29, and 0.41 per 100 patient years, respectively.
Conclusions: We describe here a workable definition for identifying major hemorrhagic events from large automated datasets. This method of identifying major bleeding may have applications for quality measurement, quality improvement, and comparative effectiveness research.

Introduction

Hemorrhage is a frequent complication of anticoagulant therapy [1]. Therefore, understanding rates of bleeding in patients receiving anticoagulation is an essential ingredient in assessing the risks and benefits of such therapy. The ability to assess rates of bleeding is important for diverse applications, including comparative effectiveness studies of different anticoagulants and efforts to measure and improve quality of anticoagulation care.

A common approach to detecting any diagnosis, especially in administrative datasets, involves using International Classification of Disease (ICD-9-CM) codes. In examining the accuracy of ICD-9-CM codes to identify complications of anticoagulation therapy, Arnason and colleagues reported a PPV of 87% for major bleeding, which increased to 96% when the bleeding code was listed as the 'most responsible diagnosis' or the 'admitting diagnosis' [2]. This study suggests that automated data alone can be sufficient to identify true-positive episodes of major hemorrhage rivaling that of chart review, particularly when additional strategies are employed to boost PPV.

The majority of previous studies assessing the accuracy of ICD-9-CM codes for identifying major hemorrhage have used chart review as a gold standard [2–6]. Additionally, the most prominent definition

Abbreviations: ARISTOTLE, Apixaban for Reduction in Stroke and Other Thromboembolic Events in Atrial Fibrillation Trial; CMS, Centers for Medicare and Medicaid Services; CPT, Current Procedural Terminology; EHR, Electronic Health Record; ICD-9-CM, International Classification of Diseases, Clinical Modification; ISTH, International Society of Thrombosis and Haemostasis; PPV, Positive Predictive Value; RELY, Randomized Evaluation of Long-Term Anticoagulation Trial; ROCKET-AF, Rivaroxaban Once daily oral direct Factor Xa inhibition Compared with vitamin K antagonist for the prevention of stroke and Embolism Trial in Atrial Fibrillation; SPORTIF, Ximelagatran Versus Warfarin for Stroke Prevention in Patients With Nonvalvular Atrial Fibrillation Trial; VA, Veterans Health Administration.

* Corresponding author at: Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, 200 Springs Road, Bedford, MA 01730. Tel.: +1 781 687 2555; fax: +1 917 591 3104.
E-mail address: guneetk@bu.edu (G.K. Jasuja).

0049-3848/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.thromres.2012.10.010

of major bleeding provided by the International Society of Thrombosis and Haemostasis (ISTH) also presupposes chart review [7]. While chart review is highly accurate, it is resource intensive. A national database of Medicare beneficiaries, for example, would require chart reviews at every hospital in the United States, which is clearly not practical. To fully harness the power of large databases to inform practice, we cannot always rely upon chart review. To our knowledge, there have been no previous attempts to adapt the ISTH definition of major hemorrhage for use with automated data.

Therefore, the objective of our study was to develop a method for identifying major hemorrhagic events in a linked Veterans Health Administration (VA)-Medicare dataset [8]. We explored the impact of different definitions and strategies upon the apparent incidence of major hemorrhage in this automated dataset and compared the rates found in our study with those reported in previous randomized and observational studies. We expected that the results of this effort would pave the way to unlocking the power of large automated datasets to track major hemorrhage for comparative effectiveness research and quality assurance/quality improvement.

Methods

Data

The data for this study included all 122,159 patients who received warfarin therapy from the VA between 10/1/06 and 9/30/08, including patients new to warfarin and those who were already experienced users. Details regarding how we built this database appear elsewhere [9]. This included demographics, ICD-9-CM diagnosis codes, and dates of service (from both sources), as well as laboratory data and pharmacy records (VA data only). This study was approved by the Institutional Review Board of the Bedford VA Medical Center.

Since we used a merged VA-Center for Medicare and Medicaid Services (CMS) dataset, we can be assured of nearly complete capture of all relevant care received by this group [8]. Out of these 122,159, 19,764 patients enrolled in Medicare Advantage program were excluded, because unlike patients participating in this capitated program would not produce itemized claims data, leading to undercounting of events. After excluding patients enrolled in Medicare Advantage, our final sample consisted of 102,395 patients who received 158,511 person-years of warfarin treatment from the VA.

For this study, we defined a period when each patient was considered to be "on warfarin" and therefore eligible to record a bleeding event. The date of warfarin inception is taken to be the first time warfarin is dispensed by the VA pharmacy or the first INR test, whichever comes first. The period begins with the latter of date of warfarin inception and 10/1/06; it ends at the latest date of a pharmacy fill plus 30-day grace period (to account for the duration of use) or an INR test, up to a maximum of 9/30/08. These 2 dates define a "window" for each patient, and only bleeding events that occur during this window can be included.

Overview of Strategy to Identify Major Hemorrhage

Our strategy was to approximate the ISTH definition for major hemorrhage as closely as possible [8], given the nature of our data. The ISTH definition uses the following criteria to define major bleeding: 1. Fatal bleeding, and/or, 2. Symptomatic bleeding in a critical area or organ such as intracranial, and/or, 3. Bleeding causing a fall in hemoglobin level of 20 g L^{-1} or more, or leading to transfusion of two or more units of whole blood or red cells [7]. For the current study, candidate bleeding codes were identified from the inpatient and outpatient ICD-9-CM codes recorded in VA and Medicare datasets. We required these codes to fulfill at least one of four criteria to be considered a major hemorrhagic event: 1) bleeding associated with death within 30 days; 2) bleeding into a critical anatomic site which necessarily would threaten life or limb; 3) bleeding associated with a transfusion of packed red blood cells or whole blood; or 4) a bleeding event characterized in our datasets as either the primary reason for a hospital admission (VA data) or the main condition (commonly known as the "principal diagnosis") for which the services are provided during a hospital stay (Medicare data).

ICD-9-CM Codes for Major Hemorrhage

We began by examining the lists of ICD-9-CM codes used by several previous studies to identify major hemorrhage, including Schalekamp et al. [5], Boulanger [6], and Arnason [2]. Thus, we began with a comprehensive list of candidate codes building on prior research. Several codes were subsequently excluded as they were felt to not be representative of major hemorrhage, e.g. 593.81, Vascular Disorders of the Kidney.

Deletion of Duplicate Events

To avoid counting multiple mentions of a single event, we developed decision rules for selecting a single record out of multiple records close in time. For example, this might occur if the course of care for a bleeding event involved 2 locations, such as a transfer from a non-VA hospital to a VA hospital. We considered records to denote a single event if they occurred on the same day or within 7 days in either direction and used the same 5-digit ICD-9-CM code. We also developed an algorithm to determine which event to retain when two conflicted; details can be found in Appendix A.

Finally, we developed an algorithm to assign a "primary" type of bleeding when codes for different types of bleeding occurred on the same date. We created five main categories of bleeding, namely intracranial bleeding, gastrointestinal bleeding, hemarthrosis, urinary bleeding, and bleeding from the throat. It was uncommon to have bleeding codes from more than one category on the same day. For patients with multiple codes within a category, we created a ranked hierarchy of codes within category to aid in selecting a single code for each episode. Details for this algorithm are found in Appendix B.

Definition: Fatal Bleeding

For the outcome of death and its date, we used the VA Vital Status Mini-File. This dataset combines multiple sources of data, including the national death index, Medicare data, and VA data to determine a single best date of death for each VA patient. The Vital Status Mini-File is considered a reliable and authoritative source for dates of death among VA patients [10].

We defined fatal hemorrhage as a bleeding event followed by death within 30 days. However, we excluded certain categories of bleeding that were implausible causes of death, including epistaxis, hematuria, bladder wall hemorrhage, hemarthrosis, and any bleeding associated with internal or external hemorrhoids. Patients who died within 30 days of such events were not considered to have died because of the bleeding.

Definition: Critical Anatomic Site

Bleeding into a critical anatomic site was defined as major bleeding, because it would necessarily threaten life or limb. This included any type of intracranial hemorrhage, hemopericardium, hemoperitoneum, and any type of hemarthrosis (which would threaten limb function). We could not include retroperitoneal hemorrhage, intraspinal, intraocular, and intramuscular bleeding with compartment syndrome, which are equally serious, because no ICD-9-CM code uniquely identifies these conditions. Additionally, to a large extent we have likely captured these relatively uncommon intraspinal and intramuscular bleeding episodes using the all-purpose ICD-9-CM code "Hemorrhage NOS" (459.0). We also separately tabulated rates for intracranial hemorrhage,

a subset of critical-site bleeding defined by the ICD-9-CM codes 430, 431, 432.0, 432.1, and 432.9.

Definition: Transfusion

We identified all transfusions of packed red blood cells or whole blood that occurred within 30 days of a bleeding event. The number of units transfused was not available. Transfusions were identified from surgical codes 99.0, 99.03, or 99.04 (Blood Transfusion, Whole Blood Transfusion, and Packed Cell Transfusion) or CPT codes 36430, 36431, or 36440 (Blood Transfusion Service).

Definition: Primary/Principal Inpatient Diagnosis Code

In VA data, the "primary" diagnosis for each inpatient admission is the ICD-9-CM code for the condition that is most responsible for the patients' length of stay. In the CMS dataset, a "principal" code was identified as the diagnosis, condition, problem or other reason for the admission/encounter/visit shown in the medical record to be chiefly responsible for the services provided. We considered any ICD-9-CM code that occupied the primary or the principal position for an inpatient stay to represent a major hemorrhage.

Statistical Analyses

We considered any candidate event which fulfilled at least one of the four criteria discussed above (fatal event, critical anatomic site, transfusion, primary reason for a hospitalization) to represent a major hemorrhage. We also performed sensitivity analyses using an alternative, more restrictive definition that omitted the criterion "primary diagnosis code". We compared, in tabular form, the proportion of events retained under each definition, stratified by ICD-9-CM code. For these analyses, we allowed multiple events and did not censor.

For age-stratified rates of bleeding, all events were classified by patient's age as of Oct. 1, 2006, not age at time of event. Confidence intervals were obtained assuming events were Poisson-distributed. All analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC). Drs. Rose and Jasuja and Mr. Reisman had full access to the study data and guarantee the accuracy and completeness of the results.

Results

Patients

Descriptive characteristics of the patients are reported in Table 1. The study sample primarily consisted of White males, with a median age of 70 years. Approximately 74% of the sample had been using warfarin for at least one year at the time of study inception. Patients in this study had a high burden of comorbid illness; for example, 31% had heart failure, 39% had diabetes, and 14% had chronic kidney disease. The burden of mental health conditions was also considerable: 23% of patients carried a diagnosis of depression, and 11% a diagnosis of alcohol abuse or dependence. Twenty-eight percent of patients were hospitalized at least once during the two-year study.

Inclusion of Candidate Events

Table 2 lists the bleeding ICD-9-CM codes for only those codes with at least 100 candidate events, and shows the number and rates for each of these bleeding event retained and the reason for their inclusion. A full version of this table, which also includes rarer event types, is presented in Supplementary Table 1 (Online Appendix). Among 71,338 unique events, 11,240 events (15.8%) were retained. Events that would seem on their face to be relatively severe were more often retained. For example, Diverticula of Colon with Hemorrhage (562.12), Esophageal Hemorrhage (530.82) and Hematemesis (578.0) were retained 46%, 43% and 42% of the time, respectively. Many candidate events, however, were contributed by event types that would only rarely have truly severe manifestations; these events were retained at much lower rates. For example, there were 30,100 episodes of Hematuria (representing 42% of candidate events), but only 4% of them qualified as major hemorrhage by our definition. Similarly, there were 5,927 episodes of Epistaxis, but only 6% were retained. Comparison of rates of inclusion and retention for our main definition and the more restrictive definition are presented in Supplementary Table 2 (Online Appendix).

Table 1
Patient characteristics (n = 102,395).

Variable	Number (%), or Median (IQR)
Female Gender	2,291 (2.2%)
Median Age (IQR)	70 (60,78)
Race/Ethnicity	
Non-Hispanic White	85,987 (84.0%)
Non-Hispanic Black	10,248 (10.0%)
Hispanic	2,821 (2.8%)
Asian	385 (0.4%)
Native American	466 (0.5%)
Other/Unknown	2,488 (2.4%)
Median Percent Poverty in Zip Code of Residence (IQR)	10.8 (6.7,16.1)
Median Distance from Nearest VA Facility in Miles (IQR)	8.0 (3.7,17.3)
Primary Indication for Warfarin	
Atrial Fibrillation	61,967 (60.5%)
Venous Thromboembolism	30,978 (30.3%)
All Others Combined	9,450 (9.2%)
Date When First Began Warfarin Therapy	
0-1 Years Prior to 10/1/06	75,622 (73.8%)
Between 10/1/06 and 9/30/07	16,576 (16.2%)
Between 10/1/07 and 9/30/08	10,197 (10.0%)
Physical Comorbid Conditions	
Cancer (newly diagnosed)	7,472 (7.3%)
Chronic Kidney Disease	13,968 (13.6%)
Chronic Liver Disease	1,350 (1.3%)
Chronic Lung Disease	29,867 (29.2%)
Diabetes	39,771 (38.8%)
Epilepsy	2,949 (2.9%)
Congestive Heart Failure	31,738 (31.0%)
Mental Health Conditions	
Alcohol Abuse	11,193 (10.9%)
Bipolar Disorder	2,723 (2.7%)
Dementia	5,110 (5.0%)
Major Depression	23,300 (22.8%)
Substance Abuse (non-Alcohol)	5,259 (5.1%)
Number of Non-Warfarin Medications	
0-7	41,946 (41.0%)
8-11	32,522 (31.8%)
12-15	17,844 (17.4%)
16+	10,083 (9.8%)
Hospitalized at least once during study	28,458 (27.8%)

Population Rates of Major Hemorrhage

We computed event rates by age strata for major hemorrhage, intracranial hemorrhage, and fatal hemorrhage for the entire population of 102,395 patients who received 158,511 patient-years of warfarin treatment from the VA (Table 3). Major hemorrhage occurred with an overall rate of 4.73 per 100 patient-years. The rate of intracranial hemorrhage, 1.29, represented 27% of all major hemorrhages, and the rate of fatal hemorrhage, 0.41, represented 9% of all major hemorrhages. There was an increasing rate of all 3 kinds of hemorrhage (any major, intracranial, and fatal) with increasing age. In particular, the rate of fatal hemorrhage increased sharply above age 70.

Comparison of the Present Study to other Major Studies

Table 4 presents a comparison of the rates of major bleeding reported in previous observational and randomized studies with the present study. The rates for major bleeding reported in observational

Table 2
Inclusion of candidate events by ICD-9-CM code. The total number of candidate events was 71,338, of which 11,240 were retained (15.8%). In this abbreviated version of the table, only event types with ≥100 candidate events are shown; the full version of the table, including rarer event types, can be found in the online appendix (Supplementary Table 1).

ICD-9 – CM code	Label	Candidate Events	Retained	Primary Diagnosis*	Transfusion*	Fatal Event*†
Bleeding into Critical Anatomic Site						
430	Subarachnoid Hemorrhage	292	ALL ‡	103 (35.3%)	6 (2.1%)	31 (9.7%)
431	Intracerebral Hemorrhage	1,317	ALL ‡	441 (33.5%)	31 (2.3%)	236 (16.1%)
432.1	Subdural Hemorrhage	1,289	ALL ‡	330 (25.6%)	29 (2.3%)	108 (7.4%)
432.9	Intracranial Hemorrhage NOS	440	ALL ‡	146 (33.2%)	7 (1.6%)	52 (11.1%)
568.81	Hemoperitoneum	221	ALL ‡	75 (33.9%)	31 (14%)	12 (5.0%)
719.16	Hemarthrosis, Lower Leg	258	ALL ‡	35 (13.6%)	8 (3.1%)	N/A
Bleeding into Non-Critical Anatomic Site						
455.2	Internal Hemorrhoids Without Complications NEC	735	68 (9.2%)	52 (76.5%)	19 (27.9%)	N/A
455.5	External Hemorrhoids With Complications NEC	166	10 (6.0%)	5 (50.0%)	5 (50.0%)	N/A
455.8	Hemorrhoids NOS With Complications NEC	232	14 (6.0%)	10 (71.4%)	4 (28.6%)	N/A
459.0	Hemorrhage NOS	1,512	231 (15.3%)	130 (56.3%)	78 (33.8%)	52 (22.5%)
530.7	Mallory-Weiss Syndrome	182	63 (34.6%)	33 (52.4%)	32 (50.8%)	4 (6.3%)
535.01	Acute Gastritis With Hemorrhage	111	42 (37.8%)	25 (59.5%)	21 (50.0%)	6 (14.3%)
535.41	Gastritis NEC With Hemorrhage	204	81 (39.7%)	39 (48.2%)	42 (51.9%)	9 (11.1%)
535.51	Gastritis/Duodenitis NOS With Hemorrhage	244	66 (27.0%)	33 (50.0%)	32 (48.5%)	4 (6.1%)
537.83	Angiodysplasia Stomach/Duodenum With Hemorrhage	178	82 (46.1%)	33 (40.2%)	54 (65.9%)	2 (2.4%)
562.12	Diverticula of Colon With Hemorrhage	552	253 (45.8%)	163 (64.4%)	99 (39.1%)	8 (3.2%)
562.13	Diverticulitis of Colon With Hemorrhage	155	62 (40.0%)	48 (77.4%)	15 (24.2%)	3 (4.8%)
569.3	Rectal And Anal Hemorrhage	5,911	616 (10.4%)	485 (78.7%)	175 (28.4%)	65 (10.6%)
569.85	Angiodysplasia with Hemorrhage NEC	197	79 (40.1%)	41 (51.9%)	41 (51.9%)	3 (3.8%)
578.0	Hematemesis	739	313 (42.3%)	220 (70.3%)	86 (27.5%)	65 (20.8%)
578.1	Blood In Stool	7,866	975 (12.4%)	725 (74.4%)	280 (28.7%)	106 (10.9%)
578.9	Hemorrhage of Gastrointestinal Tract NOS	6,837	1,898 (27.8%)	1,390 (73.2%)	583 (30.7%)	308 (16.2%)
596.7	Bladder Wall Hemorrhage	105	18 (17.1%)	9 (50.0%)	10 (55.6%)	N/A
599.7	Hematuria	30,100	1229 (4.1%)	1070 (78.3%)	366 (30.0%)	N/A
782.7	Spontaneous Ecchymoses	509	28 (5.5%)	18 (64.3%)	6 (21.4%)	5 (17.9%)
784.7	Epistaxis	5,927	346 (5.8%)	264 (76.3%)	105 (30.3%)	N/A
786.3	Hemoptysis	4,370	532 (12.2%)	326 (61.3%)	106 (19.9%)	150 (28.2%)

* These reasons for inclusion are not mutually exclusive.
† N/A for some diagnoses that were not considered a plausible direct cause of death.
‡ Bleeding into a critical anatomic site was always included as a major hemorrhage.

cohorts such as ATRIA (0.91 per 100 person-years) and ACTION (1.90 per 100 person-years) were much lower than what we found in our study. The rates for major bleeding found in randomized clinical trials ranged from 2.2 per 100 person-years (ACTIVE W, 2006) to 3.4 percent per year (ROCKET-AF, 2011). Rates of bleeding from earlier studies (as reflected in the Linkins meta-analysis) tended to be higher (the cumulative rate of all included studies was 7.2 percent per year).

Discussion

The ability to perform surveillance for adverse events is a key foundation for any program of comparative effectiveness research or quality measurement and improvement. Major hemorrhage is an adverse event common to all types of anticoagulant therapy, and therefore the ability to identify major hemorrhage in real-time, using an automated approach, would be extremely attractive. In this study, we sought to develop a comprehensive definition for identifying major hemorrhagic events among warfarin patients in a large, automated database, in a setting where chart review would not be possible.

In our population of patients receiving warfarin therapy for varied indications from the VA, we found an overall rate of major hemorrhage of 4.73 events/100 patient-years. The rates of intracranial and fatal hemorrhage were 1.29 and 0.41 events/100 patient-years, respectively. These rates of warfarin-related bleeding are similar in magnitude to the rates observed in recent large randomized trials, such as ROCKET-AF, ARISTOTLE, SPORTIF-V and RE-LY, as well as the pooled results of earlier randomized and observational studies summarized in a key meta-analysis by Linkins, et al. [1]. However, our rates for major bleeding were much higher than those reported in observational cohorts such as ACTION, ATRIA and Euro Heart Survey; this is likely attributable to a much higher illness burden among VA patients. Taken together, these results suggest that our approach to identifying major hemorrhage in warfarin patients may achieve similar results to other methods that have been employed. In addition, we note that the risk of any major hemorrhage, intracranial hemorrhage, and fatal hemorrhage among this at-risk population all increased in our study with increasing age, with a generally monotonic trend. This also echoes the findings of earlier studies [20–22] and further suggests

Table 3
Population rates of major hemorrhage and subtypes of major hemorrhage by age category, among patients receiving warfarin therapy from the VA between 10/1/06-9/30/08.

Age at Start of Window	# Patients	Aggregate Time on Therapy (patient-years)	Rate of Hemorrhagic Events (per 100 patient-years with 95% C.I.)		
			Any Major	Intracranial*	Fatal
Total	102,395	158,511	4.73 (4.62 - 4.83)	1.29 (1.23 - 1.35)	0.41 (0.38 - 0.45)
Under 55	9,865	13,958	2.29 (2.04 - 2.55)	0.63 (0.51 - 0.78)	0.10 (0.05 - 0.17)
55 – 59	14,185	21,169	2.96 (2.73 - 3.20)	1.01 (0.88 - 1.15)	0.14 (0.09 - 0.20)
60 – 64	14,436	21,882	2.94 (2.72 - 3.18)	0.85 (0.74 - 0.99)	0.19 (0.13 - 0.25)
65 – 69	11,120	17,644	3.89 (3.60 - 4.19)	1.00 (0.86 - 1.16)	0.27 (0.20 - 0.35)
70 – 74	15,001	24,105	5.26 (4.97 - 5.55)	1.49 (1.34 - 1.65)	0.42 (0.34 - 0.51)
75 – 79	15,689	25,206	6.12 (5.82 - 6.43)	1.60 (1.45 - 1.77)	0.54 (0.46 - 0.64)
80 – 84	15,095	23,895	6.70 (6.37 - 7.03)	1.80 (1.63 - 1.98)	0.72 (0.62 - 0.84)
85+	7,004	10,652	7.56 (7.04 - 8.10)	1.76 (1.51 - 2.03)	1.09 (0.90 - 1.31)

Patients may experience >1 major hemorrhage during the study.
* ICD-9-CM codes 430, 431, 432.0, 432.1, 432.9.

Table 4
Comparison of the present study to other major studies regarding rates of major hemorrhage.

Studies	Type of study	N (Patient yrs)	Major*	Intracranial*	Fatal*
ATRIA, 2003 [11]	Observational	11536 (25341)	0.91	0.46	–
ACTION, 2008 [12]	Observational	3396 (2892.1)	1.90	–	–
Euro Heart Survey, 2010 [13]	Observational	3978	1.50‡	–	–
Linkins 2003 [1]	Meta-analysis†	10757 (4374)	7.22	1.15	1.31
SPORTIF III, 2003 [14]	Randomized	3410 (4941)	1.8	–	–
SPORTIF-V, 2005 [15]	Randomized	3922 (6405)	3.1	–	–
ACTIVE W, 2006 [16]	Randomized	6706	2.2	–	0.26
RE-LY, 2009 [17]	Randomized	18,113	3.36	0.74	1.80
ROCKET-AF, 2011 [18]	Randomized	14,264	3.4‡	0.7‡	–
ARISTOTLE, 2011 [19]	Randomized	18,201	3.09‡	0.80‡	–
Present Study	**Automated Database**	**102,359 (158,511)**	**4.73**	**1.29**	**0.41**

* Rates are expressed as events per 100 person-years.
† The Linkins meta-analysis combined 29 randomized trials and 4 cohort studies, all published between 1990-2001.
‡ Rates are expressed as percent per year.

that our algorithm to identify warfarin-related major hemorrhage is working as intended.

The main limitation of this study is that we were not able to perform chart review to confirm cases of major hemorrhage, in large part because the merged VA-Medicare dataset we used would have required us to obtain data from every hospital in the United States. However, as discussed above, this study represents an important effort to find a rigorous and thoughtful approach to identifying warfarin-related major hemorrhage when a chart review is simply not possible. Our efforts borrowed heavily from previous studies that had used chart review as a gold standard [2–6]. However, no study involving chart review could have ever included a database of this size, because of the expense involved in reviewing so many charts. This study, therefore, represents an important effort to apply the results of previous (smaller) studies, which did involve chart review, to a new, larger platform, one which is ultimately necessary if automated databases are to be used for such important purposes as comparative effectiveness and quality assurance. Because Medicare data, which largely formed the basis of our study, is also the basis for so many other studies, our approach and our results may have wide applicability. However, since our at-risk population comprised of only patients on warfarin, findings from this study are only generalizable to patients on oral anticoagulant therapy such as warfarin and may not apply, for example, to patients receiving aspirin alone.

Our definition for major bleeding was actually fairly conservative, and it seems likely that we undercounted major bleeding if anything. The ISTH definition [7] includes a drop in hemoglobin level of 20 g L^{-1} to qualify an event as a major hemorrhage. We were unable to precisely characterize the time course of the hemoglobin level, particularly when using Medicare data (which does not include lab results). Therefore, hemoglobin levels were not used as a criterion for major hemorrhages in our study, and this likely resulted in considerable undercounting of events that otherwise would have qualified.

On the other hand, it is possible that our definition overcounted certain events; indeed, we cannot have avoided this entirely. In particular, our addition of a primary inpatient diagnosis code to the criteria for major hemorrhage represents an innovation, and is not contained within the original ISTH definition [7]. We considered this necessary as a way to compensate partially for our inability to establish major hemorrhage based on hemoglobin levels. In addition, this decision was supported by the finding of Arnason, et al. [2] that ICD-9 codes for hemorrhage have a 96% PPV when they are in the primary position for an inpatient stay. To us, this choice seemed like a logical way to capture at least some of the events that would have been included had we been able to track hemoglobin values. In addition, it can of course be argued that any bleeding that causes a hospitalization (or prolongs one) is, by definition, "major". Those who prefer to employ a definition more closely aligned to the original ISTH definition may wish to use our variant definition, which omitted this criterion (see Online Appendix for a comparison of the main and variant definitions). However, our main definition filtered more severe types of events to a lesser extent and filtered less severe sorts of events much more heavily. For example, "Diverticulosis of Colon with Hemorrhage" (562.13) was retained 47% of the time, while "Hematuria" (599.7) was retained only 4% of the time. In this respect, we consider our approach a success in differentiating between major and minor hemorrhagic events.

In conclusion, a definition of "warfarin-related major bleeding", modeled on similar parameters as the ISTH definition and applied to an automated dataset, allowed for computation of meaningful and comparable rates of hemorrhage. This method of extracting major hemorrhage may also prove useful with other automated datasets similar to our VA-Medicare merged dataset, for applications including comparative effectiveness, quality measurement, and quality improvement.

Conflict of Interest Statement

None.

Acknowledgement of Funding Sources

This project is supported by VA-HSR&D-IIR-10-374. Dr. Rose is supported by a VA Career Development Award (HSR&D-CDA-2-08-017). The sponsor had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; and the preparation, review, and approval of the manuscript.

Disclaimer: The opinions expressed in this manuscript do not necessarily represent the official views of the Department of Veterans Affairs.

Disclosures: Dr. Hylek has served on advisory boards for Bayer, Boehringer-Ingelheim, Bristol Myers Squibb, Daiichi Sankyo, Johnson and Johnson, Merck, and Pfizer. None of the other authors report any potential conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.thromres.2012.10.010.

References

[1] Linkins L, Choi PT, Douketis JD. Clinical impact of bleeding in patients taking oral anticoagulant therapy for venous thromboembolism. Ann Intern Med 2003;139:893-900.
[2] Arnason T, Wells PS, van Walraven C, Forster AJ. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. Thromb Res 2006;118:253-62.
[3] Andrade SE, Gurwitz JH, Chan KA, Donahue JG, Beck A, Boles M, et al. Validation of diagnoses of peptic ulcers and bleeding from administrative databases: a multi-health maintenance organization study. J Clin Epidemiol Mar 2002;55(3):310-3.
[4] Cattaruzzi C, Troncon MG, Agostinis L, García Rodríguez LA. Positive predictive value of ICD-9th codes for upper gastrointestinal bleeding and perforation in

the Sistema Informativo Sanitario Regionale database. J Clin Epidemiol Jun 1999;52(6):499-502.
[5] Schalekamp T, Klungel OH, Souverein PC, Boer A. Effect of oral antiplatelet agents on major bleeding in users of coumarins. Thromb Haemost 2008;100:1076-83.
[6] Boulanger L, Hauch O, Friedman M, Foster T, Dixon D, Wygant G, et al. Warfarin exposure and the risk of thromboembolic and major bleeding events among Medicaid patients with atrial fibrillation. Ann Pharmacother 2006;40:1024-9.
[7] Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. J Thromb Haemost 2005;3:692-4.
[8] Hynes DM, Koelling K, Stroupe K, Arnold N, Mallin K, Min-Woong Sohn, et al. Veterans' Access to and Use of Medicare and Veterans Affairs Health Care. Med Care 2007;45(3):214-23.
[9] Rose AJ, Hylek EM, Ozonoff A, Ash AS, Reisman JI, Berlowitz DR. Patient characteristics associated with oral anticoagulation control: Results of the Veterans AffaiRs Study to Improve Anticoagulation (VARIA). J Thromb Haemost Oct 2010;8(10):2182-91.
[10] Arnold N, Min-Woong Sohn, Maynard C, Hynes DM. VA-NDI Mortality Data Merge Project. VIReC Technical Report 2. Illinois: VA Information Resource Center; 2006.
[11] Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, et al. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? JAMA Nov 26 2003;290(20):2685-92.
[12] Rose AJ, Ozonoff A, Henault LE, Hylek EM. Warfarin for atrial fibrillation in community-based practise. J Thromb Haemost Oct 2008;6(10):1647-54.
[13] Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. Chest Nov 2010;138(5):1093-100.
[14] SPORTIF-III Investigators. Stroke prevention with the oral direct thrombin inhibitor ximelagatran compared with warfarin in patients with non-valvular atrial fibrillation (SPORTIF III): randomised controlled trial. Lancet 2003;362(9397):1691-8.
[15] Albers GW, Diener HC, Frison L, Grind M, Nevinson M, Partridge S, et al. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation: a randomized trial. JAMA Feb 9 2005;293(6):690-8.
[16] ACTIVE Writing Group of the ACTIVE InvestigatorsConnolly S, Pogue J, Hart R, Pfeffer M, Hohnloser S, et al. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. Lancet Jun 10 2006;367(9526):1903-12.
[17] Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med Sep 17 2009;361(12):1139-51.
[18] Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med Sep 8 2011;365(10):883-91 [Epub 2011 Aug 10].
[19] Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, et al. Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med Sep 15 2011;365(11):981-92.
[20] Hylek EM, Singer DE. Risk factors for intracranial hemorrhage in outpatients taking warfarin. Ann Intern Med 1994;120:897-902.
[21] Fang MC, Chang Y, Hylek EM, Rosand J, Greenberg SM, Go AS, et al. Advanced age, anticoagulation intensity, and risk for intracranial hemorrhage among patients taking warfarin for atrial fibrillation. Ann Intern Med Nov 16 2004;141(10):745-52.
[22] Fang MC, Go AS, Hylek EM, Chang Y, Henault LE, Jensvold NG, et al. Age and the risk of warfarin-associated hemorrhage: the anticoagulation and risk factors in atrial fibrillation study. J Am Geriatr Soc Aug 2006;54(8):1231-6.

THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE
Volume 18, Number 1, 2012, pp. 76–82

DOI: 10.1089/acm.2009.0489

Living with HIV: Responses to a Mantram Intervention Using the Critical Incident Research Method

Jeanne Kemppainen, PhD, RN,[1] Jill E. Bormann, PhD, RN,[2,3] Martha Shively, PhD, RN,[2,3] Ann Kelly, MS, APRN, BC,[2,3] Sheryl Becker, MS, RN,[2] Patricia Bone, BS, RN,[2,4] Wendy Belding, MA,[2,4] and Allen L. Gifford, MD[5]

Abstract

Objectives: The objective of this study was to identify and describe ways that a spiritually based intervention of silently repeating a mantram—sacred word or phrase—was used as a coping strategy for managing human immunodeficiency virus (HIV) disease.
Design: The design was a qualitative research method, the critical incident technique.
Settings/location: The study was conducted at an academically affiliated Veterans Affairs Hospital in southern California.
Subjects: The subjects were outpatient adults living with HIV ($n=32$) who were receiving care through HIV clinics, community agencies, and HIV providers.
Interventions: Subjects who participated in the intervention arm of a randomized controlled trial that tested the efficacy of a 5-weekly group mantram intervention were interviewed 2 months postintervention. Follow-up telephone interviews were specifically aimed at identifying instances of mantram use, and also participant perceptions of intervention usefulness or nonusefulness.
Outcome measures: The outcome measures comprised categorization and comparison of the types and frequency of incidents reported, describing ways that the intervention was "helpful" or "not helpful" in managing stressors of HIV disease.
Results: Participants reported a total of 185 incidents. Analysis and classification of the incidents resulted in eight mutually exclusive categories, including Increasing calm and/or peace, Mastering the technique, Changing my viewpoint, Increasing personal awareness, Adjusting behaviors, Managing physical symptoms, Increasing spirituality, and Enhancing relationships.
Conclusions: This study shows support for the benefits of the mantram intervention for adults with HIV. Additionally, the spiritually based mantram repetition intervention was found to be more helpful in providing a convenient, portable tool for managing a wide range of situations related to living with HIV disease.

Introduction

Although great strides have been made in the treatment of human immunodeficiency virus (HIV) disease and in extending both the quality and length of life, HIV/acquired immunodeficiency syndrome (AIDS) remains a stressful and demanding illness.[1] The psychologic and physical demands of coping with complex medication regimens, side-effects of treatment, symptom severity, as well as comorbid illnesses can be overwhelming and may influence health care behaviors such as medication adherence, substance use, sexual risk taking, or other behaviors that influence health outcomes.[1,2] In addition, persons living with HIV disease often experience psychosocial challenges such as depression, isolation, and traumatic life events that can impact the course of HIV disease progression.[2,3] Evidence suggests that stress may hasten HIV disease progression by increasing viral replication and suppressing the immune response.[4–6]

Intervention studies using combinations of cognitive and relaxation therapies in the era of highly active antiretroviral

[1]University of North Carolina Wilmington, Wilmington, NC.
[2]VA San Diego Healthcare System, San Diego, CA.
[3]Nursing Service for Research, VA San Diego Healthcare System, San Diego, CA.
[4]Veterans Medical Research Foundation, San Diego, CA.
[5]Center for Health Quality, Outcomes & Economic Research, Boston University School of Public Health, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA.

therapy have shown improvements in psychologic quantitative measures in HIV-infected persons.[7] Limitations of these studies are that (1) they assessed multiple-component interventions (e.g., a variety of relaxation techniques and cognitive–behavioral strategies), making it difficult to conclude which ones accounted for changes in health status; (2) most were conducted on samples of gay and bisexual men, thus failing to study women and heterosexuals; and (3) most were pre-experimental or had a wait-list control group rather than a comparable control group.

Although research findings are not entirely consistent, there is some evidence that spirituality predicts slower HIV disease progression.[8,9] Patients who have endorsed a greater sense of spirituality after being diagnosed with HIV have shown a slower decline in CD4 cell counts and better control of viral load over 4 years.[8]

Recent attention has been directed toward developing spirituality-based interventions for managing stressors of chronic illness, including HIV/AIDS.[10–12] Practices such as prayer and meditation, for example, have been cited as commonly used coping strategies by persons living with HIV/AIDS.[13–15] More research on the healing aspects of spirituality is needed as new methods to measure spirituality are becoming refined.[16–19]

Objectives

In light of the huge impact of the psychologic burden experienced by persons living with HIV/AIDS, it is important to gain a broader understanding of the specific uses that a mantram-based, spiritual intervention has on the perceptions and behavioral actions of persons living with HIV disease. To provide a more comprehensive picture of the usefulness of a mantram intervention developed by Bormann and colleagues,[20] a qualitative component was added to the larger randomized controlled trial (RCT). Follow-up interviews were conducted with participants in the mantram arm of the original study at 2 months postintervention to enhance the quality and credibility of the quantitative findings.[21] The interviews were specifically aimed at determining how the skills learned from a mantram intervention played out in the real life of study participants. This analysis examined the qualitative interview findings. Specific objectives for this analysis included (1) identifying incidences of mantram use following participation in a spiritually based mantram intervention, and (2) identifying participant perceptions of intervention usefulness or nonusefulness.

Materials and Methods

Procedures

In the original RCT, Bormann and colleagues examined the efficacy of a group-based mantram intervention on HIV outcomes with a sample of HIV-infected adults receiving outpatient care through HIV clinics, community agencies, and HIV providers.[20] The study was conducted through the VA San Diego Health Care System and the University of California San Diego. Human subject approvals were obtained from university and VA hospital committees.

In the original RCT, participants were randomly assigned to either a mantram-based intervention or to an attention-matched education control group. Participants attended a series of 5 weekly 90-minute sessions, followed by 4 weekly automated phone calls from co-facilitators, and a final booster session in week 10. Sessions consisted of how to (1) choose and use a mantram, (2) practice slowing down one's thinking to help set health-related priorities, and (3) develop one-pointed attention for stress management. Participants were given *The Mantram Handbook*,[22] a list of mantrams to choose from and course manual with homework exercises. Weekly assignments included practicing mantram repetition at non-stressful times, such as each night before sleep or while waiting in lines. Other strategies taught were slowing down mentally and behaviorally, to make wiser choices, set priorities, and decrease hurried behavior, and one-pointed attention to increase concentration for repeating the mantram or engaging in one task at a time. For a more complete description of the intervention, see Bormann, 2006.[20] The qualitative data in this analysis were obtained through follow-up telephone interviews conducted at 2 months postintervention to assess the uses and sustainability of mantram practice.

Based on the findings of a larger RCT, it was anticipated that this spiritually based intervention would foster the continued use of positive coping behaviors in persons with HIV/AIDS.

Interviews

A qualitative research method, the critical incident technique,[23] was chosen for its ability to pinpoint specific behaviors and benefits of the mantram. This valid and reliable interviewing format is a very practical and efficient method for obtaining information quickly and in the participants' own words.[24] Critical incident interviews may be very brief, perhaps as short as 10–15 minutes. They are designed to pinpoint facts and eliminate personal opinion or generalizations.[23] They require only simple types of judgments and responses from the participants, and are used to increase knowledge about little-known phenomena, such as mantram repetition.

The critical incident technique has been applied in a variety of health care settings to examine patients' met and unmet psychologic needs, [25] predict patient perceptions of nurse behaviors,[26] and gather information about extended care nursing[27] and long-term care facilities.[28]

In the current study, critical incident interviews were conducted as follows. Participants were contacted 2 months after the final group session by a research nurse, trained in the critical incident interviewing method. Telephone audiotaped interviews were conducted on all participants, asking them to recall specific incidents of mantram use. Participants were asked to describe as many situations as could be remembered and to give examples of successful outcomes, as they defined it. Standard probes were used to obtain a more complete description of the context of the incident. The critical incident interviews lasted approximately 15 minutes and were transcribed for analysis. A critical incident expert was consulted prior to and during the study to guide the methods and analysis (e.g., taught the interviewer how to ask questions and assisted with data interpretation).

Data preparation and analysis

The data obtained through the critical incident interviews were analyzed through an inductive classification process developed by Flanagan.[23] This careful process enabled the

researchers to build a comprehensive picture of the behavioral dimensions of mantram use.[29]

First, the transcribed interview data were audited to ensure that each of the incidents included complete behavioral descriptions and also demonstrated a linkage between what the respondent did or did not do and the resulting outcome. The members of the research team and a doctorally prepared nurse with expertise in the critical incident research method made these judgments. Only 185 incidents out of 200 or 93% met Flanagan's criteria and were included in the analysis.

Incidents, which were judged to be nearly identical or very similar, were grouped together. Similar groupings were then combined to form subcategories of behaviors. Subcategories were sorted and grouped together to define more inclusive major categories of mantram use. Members of the research team and two doctoral-prepared nurses with expertise in HIV/AIDS and research methods independently re-sorted the incidents, and their disagreements were used to refine and determine the final set of eight mutually exclusive and exhaustive major behavioral categories.

As a check on the comprehensiveness of the classification system, 10% of the incidents were randomly selected using a table of random numbers and were independently categorized by 2 expert reviewers. The percent of inter-rater agreement was 95%.

Results

Subjects

Table 1 shows the demographic and severity of illness characteristics for the participants included in this analysis ($n=32$). The majority of the sample was male (78.1%) with a mean age of 44.1 years (standard deviation [SD] 6.7). Forty-three percent (43%) of the participants were nonwhite. The mean CD4 count for the sample was 517/mm^3 (SD 280.43), with an average time since acquiring an HIV diagnosis of 8.3 years (SD 5.96). At the time of the study, 23 of 32 participants (71%) were receiving highly active antiretroviral therapy. The majority of the respondents reported that they were never married or partnered (59.4%; $n=19$), while 81% ($n=70$) indicated that they completed some college. Sixty-nine percent (69%) ($n=22$) identified with a religious group. Approximately 41% of participants scored above the cutoff for depressive symptoms on the Clinical Epidemiology Study-Depression scale.

Table 1. Demographic Data ($N=32$)

Variable	*Range*	*Mean (yrs)*	*SD*	*Frequency*
Age	29–57	44.1	6.7	32
Years with HIV	0.5–25	8.3	5.96	32
Sex				
Male				25 (78%)
Female				7 (22%)
Ethnicity				
White				18 (56.3%)
Black				11 (34.4%)
Hispanic				3 (9.4%)
Risk factors				
Male/male sex				18 (27.3%)
Heterosexual				11 (16.7%)
Don't know				8 (12.1%)
IV drug use				5 (7.6%)
Transfusion recipient				1 (1.5%)
Education				
High school or less				8 (25.0%)
Some college				18 (56.3%)
College degree or higher				6 (18.8%)
Marital status				
Never married/ partnered				19 (59.4%)
Ever been married/ partnered				13 (40.6%)
Employment				
>20 hrs/wk				14 (43.8%)
≤20 hrs/wk				9 (28.1%)
None				9 (28.1%)
CD4 cell count				
≥200				27 (84.4%)
<200				5 (15.6%)
HIV-PCR				
<400				21 (65.6%)
401–30,000				7 (21.9%)
30,000–100,000				3 (9.4%)
>100,000				1 (3.1%)
Receiving HAART				
Yes				23 (71.9%)
No				9 (28.1%)
Depression (CES-D)				
None				13 (40.6%)
Mild to moderate				6 (18.8%)
Major				13 (40.6%)
Identifies with a religious group				
Yes				22 (68.8%)
No				10 (31.3%)

SD, standard deviation; IV, intravenous; HIV-PCR, human immunodeficiency virus–polymerase chain reaction; HAART, highly active anti-retroviral therapy; CES-D, Center for Epidemiological Studies Depression Scale.

Critical incident analysis

A total of 185 incidents describing mantram use behaviors were obtained from 32 participants who completed the mantram intervention. The analysis and classification of the incidents resulted in 8 mutually exclusive categories, and 16 subcategories.

Major categories of mantram use included the following: *Increasing calm and/or peace, Mastering the technique, Changing my viewpoint, Increasing personal awareness, Adjusting behaviors, Managing physical symptoms, Increasing spirituality, and Enhancing relationships.* A review of the incidents that were sorted into each category determined the definition. The taxonomy of major mantram repetition behaviors is presented in Table 2.

The largest category, *increasing calm and/or peace*, accounted for 36% of the total number of incidents. Two thirds of the incidents placed within this category ($n=42$; 64%) described behaviors where mantram repetition increased a sense of peace or calm in response to an identified stressor, including waiting for a doctor appointment, encountering a stressful situation while riding on public transportation, or being late for an appointment because "someone was late picking me up." As one of the participants stated:

Table 2. Categories of Mantram Repetition Behaviors (N=185 Incidents)

Major category	*Number of incidents*	*Percent*
I. Increasing calm and/or peace (66 incidents; 36%)		
1. Increasing calm and/or peace with an identified stressor	42	64%
2. Increasing calm and/or peace *without* an identified stressor	11	17%
3. Replacing anger with peace	11	17%
4. Calming others around me	2	3%
II. Mastering the technique (39 incidents; 21%)		
1. Practicing Mantram repetition	20	51%
2. Using one-pointed attention	16	41%
3. Renewing previous mantram training	3	8%
III. Changing my viewpoint (36 incidents; 19%)		
1. Changing perspective and "letting go of concern"	23	64%
2. Increasing patience	7	19%
3. Controlling thoughts	6	17%
IV. Increasing personal awareness (15 incidents; 8%)		
1. Becoming aware of emotional states	12	80%
1. Becoming aware of physical state	3	20%
V. Adjusting behaviors (14 incidents; 7.6%)		
1. Changing a specific behavior	12	86%
2. Slowing down	2	14%
VI. Managing symptoms (10 incidents; 5%)		
1. Inducing sleep	8	80%
2. Finding pain relief	1	10%
3. Breathing easier	1	10%
VII. Increasing spirituality (4 incidents; 2%)	4	100%
VIII. Enhancing relationships (1 incidents; 0.05%)	1	100%

> When I am really frustrated or in a line or something, I don't let that bother me. I just say my mantram, and before you know it, I am right up at the front of the line. It has really worked for me.

Mantram repetition was also used for managing emotions related to driving in heavy traffic or encountering "road rage." Many participants found that briefly repeating a mantram generated a sense of calm, and reduced negative feelings toward other drivers. Although Easwaran[22] did not endorse using mantram while driving, mantram repetition appeared to be more helpful than harmful. No one reported any accidents attributed to mantram repetition while driving. The following example describes a traffic-related incident:

> Instead of shouting or yelling or swearing at somebody in traffic, I just start using my mantram to bring myself off that cliff. I use it just about every time…. Instead of throwing words out the window, I can just sort of say it to myself….I don't drive away from the situation with my blood boiling, all mad and upset, because I realize that I keep myself down and don't let myself get to that hyped-up point.

Additional incidents placed into this category involved *increasing calm without an identified stressor* (n=11; 17%). As one participant stated:

> Sometimes when I get stressed out....like if I was having a really, really bad day…I deal with the anxiety and stress by using the mantram.

An additional 17% of the incidents in this category described the use of mantram repetition to replace anger with a sense of peace or calm. As one participant stated:

> The mantram, in itself, was very helpful. It helped out in a lot of different situations…being confronted by rude people, like in the public, and to keep from promoting violence, you know, just to back off and repeat it.

Two (2) additional incidents grouped within this major category describe the effectiveness of mantram repetition in calming others.

The second category, *mastering the technique,* reflected incidents where subjects practiced the primary skills of the mantram course intervention, including practicing mantram repetition (n=20; 51%), using one-pointed attention (n=16; 41%), and renewing previous mantram training (n=3; 8%).

The third category, *changing my viewpoint*, accounted for 16% of the total number of incidents. The majority of the incidents within this category (n=23; 77%) described circumstances where the study participants changed perspective about a circumstance or event. As 1 participant stated:

> I used to stress out about going to the doctor but don't do that anymore. I say my mantram instead. I had a colon test that I put off for a year and a half. I was so afraid thinking, "Oh my God, what if I have colon cancer?" And I scared myself so bad. I thought I'm going to get this done. And I called the doctor and I scheduled it. In the past, I would have been such a drama person about this and afraid that I had something before I went in.…I changed my whole thing of having a negative thing and going into a positive and not putting that much worth on it.

The fourth category, *increasing personal awareness,* included 15 (8%) incidents that described the linkages between mantram use and an increasing awareness of emotional and physical states. As 1 participant stated, "It taught me to feel how different my brain felt when I was calm so then I could sense that when I was getting angry."

Incidents that described behavior changes related to mantram repetition were placed in the fifth category. The following examples of changing behaviors include:

> When I get ready to eat, I turn the TV off and I just eat and just enjoy my meal. Instead of like eating and watching TV or listening to music, now I just relax and take time for myself.
>
> I am trying to be more patient with people that are not as quick as I need them to be at the moment or when I am talking to someone who is a little slower in understanding. I'm trying to also stop interrupting people and finish things for them.

A sixth category contained incidents that described the impact of mantram repetition on managing physical symptoms. Within this category, eight incidents described the use of mantram repetition as a sleep aid:

> It helps me to use the mantram instead of just lying in bed and tossing and turning. I can use my mantram and just repeat my mantram and I fall back asleep.…The next thing you know, it's the next morning.

One (1) participant described the successful use of mantram repetition to "ease breathing" during an episode of *Pneumocystis carinii* pneumonia (PCP), while a second participant reported using mantram repetition to ease physical pain. The two remaining categories included *increasing spirituality* ($n = 4$, 2%) and *enhancing relationships* ($n = 1$).

Perceptions of usefulness or nonusefulness of mantram program

Of the total number of reported incidents, the majority (99%) were positive reports of instances when mantram repetition enhanced the respondent's coping. Two (2) negative incidents associated with mantram use were reported, including the following:

> One of the things I was presented with was single task activity. I can't say that it was helpful and it was certainly contrary to everything I have ever done. I find it difficult to utilize that concept and I don't see how it can be used effectively.
> I learned the futility of trying to document the times and numbers of the mantra use. I thought it was distracting and counterproductive."

Discussion

This critical incident study represents a comprehensive effort to identify specific patterns of mantram use by persons living with HIV disease following participation in a spiritually based education program. This practical and efficient qualitative method allowed the researchers to gain a greater depth of understanding about the perceptions of study participants, as well as precise descriptions of behavioral outcomes of the mantram intervention. Participants reported using mantram repetition up to 2 months following the intervention, suggesting the sustainability of the intervention. An additional important finding is that the major categories of mantram use reflected the content of the educational program designed to teach this competency, as well as the findings of the larger RCT.

Eight (8) major categories of mantram use behaviors emerged from the critical incident interviews. The major categories from most to least frequent number of incidents included the following: *Increasing calm and/or peace, Mastering the technique, Changing my viewpoint, Increasing personal awareness, Adjusting behaviors, Managing physical symptoms, Increasing spirituality, and Enhancing relationships.* The fact that these major categories of mantram use were made up of multiple, similar behaviors lends credibility to their reports.[23,29] In addition, the reliability checks and the rigor of the critical incident technique suggest that the categories also have a consistent meaning to the study participants, as well as to the members of the research team who derived them.[29]

The largest category of mantram repetition behaviors, *Finding calm or peace,* reflected some findings that were not as clearly identified in the larger RCT. For example, results from the critical incident analysis revealed that study participants used mantram repetition to manage specific stressful situations and to decrease feelings of anxiety. In the quantitative measures, however, there were decreases of anxiety and perceived stress over time, but these results were found in both mantram and control groups, and could not be attributed to only the intervention. On the other hand, the qualitative critical incident data indicated that participants used mantram repetition to manage emotions other than stress, including anger and frustration. These findings are similar to the larger RCT,[20,30] where significant improvements were found in reducing anger, and these reductions were mediated by increased appraisal coping.

The second most frequent category of mantram repetition behaviors, *mastering the technique,* described incidents where study participants continued to practice strategies taught during the intervention. This finding further expands the quantitative outcome results of Bormann and colleagues,[20] who found significant relationships between frequency of mantram practice using wrist counters and quality of life, spiritual well-being, and subscales of meaning/purpose and faith/assurance. Although the frequency of mantram practice may have had a valuable, therapeutic effect on psychologic distress, this interpretation must be viewed with caution because participants who practiced mantram more frequently were self-selected.

The majority of the remaining categories directly reflect how the group-based mantram intervention was designed to promote increased metacognitive awareness to stressors as well as to strategies for evaluating cognitive and emotional responses to stressors. In the third most frequent category, *Changing my viewpoint,* participants reported that the use of mantram repetition led to changes in their primary appraisals of stressful events, as well as changes in their appraisals of methods for coping with those events. The critical incidents in the fourth most frequent behavioral category, *Increasing personal awareness,* indicated that mantram repetition cultivated increased metacognitive skills and heightened awareness of stressful events and to their own internal responses to stressors. A fifth category, *Adjusting behaviors,* reflected changes in their coping skills. Linkages were noted between the major categories, *Changing my viewpoint, Increasing personal awareness*, and *Adjusting behaviors* and the quantitative study findings from the RCT. In a secondary analysis performed on quantitative data from the original RCT, Bormann and Carrico[30] found significant increases in positive reappraisal coping in the mantram arm participants over the 5-week intervention period.

Analysis of the critical incident data also reflects consistency between the eighth category, *Increasing spirituality,* and the quantitative findings of the larger RCT. Bormann and colleagues[20,31] found that participants in the mantram arm of the RCT increased both spiritual faith/assurance and spiritual connectedness as well as significant correlations between frequency of mantram practice sessions and improvements in spiritual well-being subscales (i.e., meaning/peace and faith/assurance).

Although linkages were found between the many behavioral categories and the quantitative findings of the larger RCT, a difference was noted for the incidents in category six, *Managing symptoms.* A clear connection between the incidents in this category and the quantitative data in the larger RCT was not evident. The influence of spirituality on the management of physical symptoms, however, has received increased attention in the literature.[32–34] Wachholtz and Pargament[35] found spiritual meditation was related to greater decreases in both migraine headache frequency and severity compared to secular meditation. Coleman and colleagues[13,36] found that prayer was used as a self-care strategy for managing HIV-related symptoms, including depression, fatigue, and nausea. Wolf and Abell[37] found that repetition of the mantra compared to a placebo mantra or no-treatment control resulted in decreased depression and anxiety in adults.

Conclusions

Study findings support the integration of qualitative methods within a RCT to expand quantitative outcome results, and also to gain a broader understanding of how an intervention plays out in the real life of the study participants. In this study, mantram repetition was identified as a coping strategy for managing stressors associated with HIV disease. Participants applied mantram repetition to their lives, targeting areas of most bothersome distress. Study findings also contribute to a further refinement of the spiritually based intervention with veterans, and lay a foundation for use with broader populations.

Acknowledgments

This research was funded by the National Center of Complementary/Alternative Medicine, National Institutes of Health (NCAAM/NIH) #R21 AT01159 and supported by resources from the Department of Veterans Affairs, VA San Diego Healthcare System, and the Veterans Medical Research Foundation. The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs, United States Government, or NCCAM/NIH.

Disclosure Statement

No financial conflicts exist.

References

1. Moskowitz J, Hult J, Bussolari C, Acree M. What works in coping with HIV: A meta analysis with implications for coping with serious illness. Psychol Bull 2009;135:121–141.
2. Gore-Felton C, Koopman C. Behavioral mediation of the relationship between psychosocial factors and HIV disease progression. Psychosom Med 2008;70:569–574.
3. Berg C, Michaelson S, Safren S. Behavioral aspects of HIV care: Adherence, depression, substance use, and HIV transmission behaviors. Infect Dis Clin North Am 2007;21:191–200.
4. Balbin EG, Ironson GH, Solomon GF. Stress and coping: The psychoneuroimmunology of HIV/AIDS. Bailliere's Clin Endocrinol Metab 1999;13:615–633.
5. Robinson FP, Mathews HL, Witek-Janusek L. Stress reduction and HIV disease: A review of intervention studies using a psychoneuroimmunology framework. J Assoc Nurses AIDS Care 2000;11:87–96.
6. Evans DL, Leserman J, Perkins DO, et al. Severe life stress as a predictor of early disease progression in HIV infection. Am J Psychiatry 1997;154:630–634.
7. Scott-Sheldon LA, Kalichman SC, Carey MP, Fielder RL. Stress management interventions for HIV+ adults: A meta-analysis of randomized controlled trials, 1989 to 2006. Health Psychol 2008;27:129–139.
8. Ironson G, H'Sien H. Do positive psychosocial factors predict disease progression in HIV-1? A review of the evidence. Psychosom Med 2008;70:546–554.
9. Fitzpatrick AL, Standish LJ, Berger J, et al. Survival in HIV-1 positive adults practicing psychological or spiritual activities for one year. Altern Ther Health Med 2007;13:18–24.
10. Margolin A, Beitel M, Schuman-Oliver Z, Avants SK. A controlled study of a spiritually-focused intervention for increasing motivation for HIV prevention among drug users. AIDS Educ Prev 2006;18:311–322.
11. Tarakeshwar N, Pearce MJ, Sikkema KJ. Development and implementation of a spiritual coping group intervention for adults living with HIV/AIDS: A pilot study. Mental Health Religion Culture 2005;8:179–190.
12. Tuck I, McCain NL, Elswick RKJ. Spirituality and psychosocial factors in persons living with HIV. J Adv Nurs 2001; 33:776–783.
13. Coleman C, Eller L, Nokes K, et al. Prayer as a complementary health strategy for managing HIV-related symptoms among ethnically diverse patients. Holistic Nurs Pract 2006;20:65–72.
14. Cotton S, Puchalski CM, Sherman SN, et al. Spirituality and religion in patients with HIV/AIDS. J Gen Intern Med 2006; 21S5:S5–S13.
15. Duggan J, Peterson WS, Schutz M, et al. Use of complementary and alternative therapies in HIV-infected patients. AIDS Patient Care STDS 2001;15:159–167.
16. Fetzer Institute & National Institute of Aging Working Group. Multidimensional measurement of religiousness/spirituality for use in health research: A report of the Fetzer Institute/National Institute on Aging Working Group. Kalamazoo, MI: Fetzer Institute & National Institute on Aging Working Group, 2003.
17. Ironson G, Solomon GF, Balbin EG, et al. The Ironson-Woods spirituality/religiousness index is associated with long survival, health behaviors, less distress, and low cortisol in people with HIV/AIDS. Ann Behav Med 2002;24:34–48.
18. Life Advance, I. The Spiritual Well-Being Scale. Online document at: www.lifeadvance.com Accessed February 3, 2003.
19. Peterman AH, Fitchett G, Brady MJ, et al. Measuring spiritual wellbeing in people with cancer: The Functional Assessment of Chronic Illness Therapy–Spiritual Well-Being Scale (FACIT-Sp). Ann Behav Med 2002;24:49–58.
20. Bormann JE, Gifford AL, Shively M, et al. Effects of spiritual mantram repetition on HIV outcomes: A randomized controlled trial. J Behav Med 2006;29:359–376.
21. Vuckovic N. Integrating qualitative methods in randomized controlled trials: The experience of the Oregon Center for Complementary and Alternative Medicine. J Altern Complement Med 2002;8:225–227.
22. Easwaran E. The Mantram Handbook, 4th ed. Tomales, CA: Nilgiri Press, 2001.
23. Flanagan J. The critical incident technique. Psychol Bull 1954;51:327–358.

24. Kemppainen JK. The critical incident technique and nursing care quality research. J Adv Nurs 2000;32:1264–1271.
25. Kent G, Wills G, Faulkner A, et al. Patients reactions to met and unmet psychological need: A critical incident analysis. Patient Educ Counsel 1996;28:187–190.
26. Kemppainen JK, O'Brien L, Corpuz B. The behavior of AIDS patients toward their nurses. Int J Nurs Studies 1998;35: 330–338.
27. Cheek J, O'Brian B, Ballantyne A, Pincombe J. Using critical incident technique to inform aged and extended care nursing. Western J Nurs Res 1997;18:667–682.
28. Grant NK, Reimer M, Bannatyne J. Indicators of quality in long-term care facilities. Int J Nurs Studies 1996;33:469–478.
29. NcNabb W, Wilson-Pessano S, Jacobs A. Critical self-management competencies for children with asthma. J Pediatr Psychol 1986;11:103–117.
30. Bormann JE, Carrico AW. Increases in positive reappraisal coping during a group-based mantram intervention mediate sustained reductions in anger in HIV-positive persons. Int J Behav Med 2009;16:74–80.
31. Bormann JE, Smith TL, Shively M. et al. Self-monitoring of a stress reduction technique using wrist-worn counters. J Healthcare Qual 2007;29:47–55.
32. Paloutzian RF, Park CL, eds. Handbook of the Psychology of Religion and spirituality. New York: Guilford Press, 2005.
33. Plante TG, Thoresen CE. Spirit, Science and Health: How the Spiritual Mind Fuels Physical Wellness. Westport, CT: Praeger, 2007.
34. Wachholtz A, Pearce M, Koenig H. Exploring the relationship between spirituality, coping, and pain. J Behav Med 2007;30:311–318.
35. Wachholtz A, Pargament KI. Migraines and meditation: Does spirituality matter? J Behav Med 2008;31:351–366.
36. Coleman C, Holzemer W, Eller L, et al. Gender differences in use of prayer as a self-care strategy for managing symptoms in African Americans living with HIV/AIDS. J Assoc Nurs AIDS Care 2006;17:16–23.
37. Wolf DB, Abell N. Examining the effects of meditation techniques on psychosocial functioning. Res Social Work Pract 2003;13:27–42.

Address correspondence to:
Jeanne Kemppainen, PhD, RN
University of North Carolina Wilmington
601 South College Road
Wilmington, NC 28403

E-mail: kemppainenj@uncw.edu

Journal of Substance Abuse Treatment 43 (2012) 389–396

Contents lists available at SciVerse ScienceDirect

Journal of Substance Abuse Treatment

Effect of quality chronic disease management for alcohol and drug dependence on addiction outcomes

Theresa W. Kim M.D. [a,*], Richard Saitz M.D., M.P.H. [a,b], Debbie M. Cheng Sc.D. [a,c], Michael R. Winter M.P.H. [d], Julie Witas M.S.W. [a], Jeffrey H. Samet M.D., M.A., M.P.H. [a,e]

[a] Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center, Boston University School of Medicine, Boston, MA, USA
[b] Department of Epidemiology, Boston University School of Public Health, Boston, MA 02118, USA
[c] Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA
[d] Department of Data Coordinating Center, Boston University School of Public Health, Boston, MA 02118, USA
[e] Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA 02118, USA

ARTICLE INFO

Article history:
Received 11 March 2012
Received in revised form 7 June 2012
Accepted 12 June 2012

Keywords:
Primary care
Chronic disease management
Integrated care
Quality of care

ABSTRACT

We examined the effect of the quality of primary care-based chronic disease management (CDM) for alcohol and/or other drug (AOD) dependence on addiction outcomes. We assessed quality using (1) a visit frequency based measure and (2) a self-reported assessment measuring alignment with the chronic care model. The visit frequency based measure had no significant association with addiction outcomes. The self-reported measure of care—when care was at a CDM clinic—was associated with lower drug addiction severity. The self-reported assessment of care from any healthcare source (CDM clinic or elsewhere) was associated with lower alcohol addiction severity and abstinence. These findings suggest that high quality CDM for AOD dependence may improve addiction outcomes. Quality measures based upon alignment with the chronic care model may better capture features of effective CDM care than a visit frequency measure.

1. Introduction

Although alcohol and other drug (AOD) dependence is characterized as a chronic disease (McLellan, Lewis, & O'Brien, 2000), its treatment in the United States is too often episodic, poorly coordinated, and difficult to access (Friedmann, Lemon, Stein, & D'Aunno, 2003; Institute of Medicine, 2006). Access issues are due to several reasons including lack of insurance, inadequate supply of available treatments in close proximity, and confusing or strict program entry requirements (Cohen, Feinn, Arias, & Kranzler, 2007; Institute of Medicine, 2006). In addition, most patients with addictions have medical and/or psychiatric comorbidity (de Alba, Samet, & Saitz, 2004; Grant et al., 2004; Mertens, Lu, Parthasarathy, Moore, & Weisner, 2003). Care for conditions in these three spheres often occurs in separate locations and systems, a fact likely responsible in part for poor quality care (Samet, Friedmann, & Saitz, 2001).

There is growing interest in shifting addiction treatment from an "acute care framework" to one that is more suitable for patients with chronic disease (Institute of Medicine, 2006; McKay, 2009). Chronic disease management (CDM) models based on the principles of the chronic care model (Wagner, Austin, & Von Korff, 1996) may improve the quality of addiction care (Institute of Medicine, 2006; Saitz, Larson, LaBelle, Richardson, & Samet, 2008; Watkins, Pincus, Tanielian, & Klein, 2003). As with other chronic diseases such as depression, diabetes, and congestive heart failure (Blonde, 2000; Gilbody, Bower, Fletcher, Richards, & Sutton, 2006; McAlister, Lawson, Teo, & Armstrong, 2001; Neumeyer-Gromen, Lampert, Stark, & Kallischnigg, 2004), CDM for AOD dependence has the potential to improve addiction outcomes by providing patient-centered, longitudinal care to increase the receipt of effective treatments and provide self-management support (Wagner et al., 2001). Care may be delivered in a flexible manner with treatment intensity and modality responsive to fluctuations in disease severity and other patient needs. A multidisciplinary team can track and coordinate care with the goals of addressing a comprehensive set of needs (addiction, medical, psychiatric, and social) and re-engaging patients who drop out of care. However, despite the potential benefits of CDM for AOD dependence, few data are available about its effectiveness for people with AOD.

Interest in CDM for AOD dependence parallels a realignment of primary care settings to deliver longitudinal, coordinated care for a comprehensive set of health needs, known as the patient centered medical home (PCMH) (American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, American Osteopathic Association, 2011). Increasingly, policy makers and others are recognizing the importance of addressing AOD dependence in primary care to potentially realize the benefits of the PCMH (Croghn & Brown, 2010). The National Committee for Quality

* Corresponding author. Boston Medical Center, 801 Massachusetts Avenue, Boston, MA 02118, USA. Tel.: +1 617 414 6932, fax: +1 617 414 4676.
E-mail address: theresa.kim@bmc.org (T.W. Kim).

0740-5472/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsat.2012.06.001

Assurance (NCQA) recently added a PCMH accreditation standard to target a mental health, behavioral, or substance abuse "condition" for quality improvement (NCQA, 2011). Although there is growing support to integrate care for AOD dependence in primary care (Substance Abuse and Mental Health Services Administration, 2012), it is not clear how AOD dependence care should be organized and delivered in the PCMH (Croghn & Brown, 2010). CDM for AOD shares many of the core tenets of the PCMH and may be an effective strategy for addressing AOD dependence.

Although there is evidence suggesting the efficacy of CDM for AOD dependence, little is known about whether the quality of CDM care should be considered, and if so, how it should be measured. The quality of addiction care has been measured by visit frequency and timing-based measures. The quality of CDM has been measured by self-reported validated scales. How the quality of CDM relates to addiction outcomes is unknown; however, we do know that when offered, most patients with AOD dependence are willing to engage with CDM care (Kim et al., 2011).

Therefore, the objective of the current study is to examine if higher quality CDM for AOD dependence is associated with better addiction outcomes (abstinence, addiction severity). We hypothesized that receipt of quality CDM care for AOD dependence is associated with abstinence and lower addiction severity which is the primary aim of this analysis. Because CDM care is structured to increase the receipt of effective treatments, the secondary objectives are to examine whether quality CDM care is associated with higher odds of specialty addiction treatment utilization, addiction pharmacotherapy, and mutual help group attendance. The rationale for examining the latter is that mutual help group attendance is consistent with two chronic care model principles, namely the importance of helping patients develop skills and confidence to manage their addiction (i.e., a self-management plan) and of promoting involvement with community support. Mutual help groups are also often thought to be important components of addiction treatment plans in the United States.

2. Methods

2.1. Study design

This is a secondary data analysis of data collected during the Addiction Health Evaluation and Disease management (AHEAD) study, a randomized clinical trial of access to a primary care-based CDM clinic for AOD dependence. Details of the study have been published (Kim et al., 2011). Briefly, eligible participants were adults with AOD dependence (Composite International Diagnostic Interview Short Form; Kessler, Andrews, Mroczek, Ustun, & Wittchen, 1998) and any recent (past 30-days) opioid, stimulant, or heavy alcohol use (i.e., >21 drinks in a week or at least 2 days of ≥5 drinks in a day for men; >14 drinks per week or at least 2 days of drinking ≥4 for drinks in a day for women) who were willing to establish or continue primary care at Boston Medical Center. Interest in addiction treatment and readiness to change were not required for study enrollment.

Recruitment occurred in several ways: screening in an inpatient detoxification unit, referrals from Boston Medical Center mostly from primary care clinics and the emergency department, and advertising to the general public in medical center public areas, in newspapers and on buses.

After completion of the baseline research interview, subjects were randomized to either have access to the AHEAD study the CDM clinic established for that purpose or usual care. All subjects, both intervention and controls, were referred to primary care at Boston Medical Center and were given access to short-term motivational enhancement therapy (MET). All subjects could choose to access medical, psychiatric, and addiction treatment services provided by the hospital or in the community.

For subjects randomized to attend the CDM clinic, research associates accompanied them to their first visit. Subjects were compensated for study procedures completed at study entry, which included the first visit to the CDM clinic for subjects randomized to it. Beyond the initial CDM clinic visit, subjects were neither required to attend nor compensated for additional CDM clinic visits though that clinic remained available to them. In-person research interviews were done at 3, 6, and 12 months after study enrollment. Research associates neither encouraged nor discouraged CDM clinic use.

2.2. Description of the AHEAD study CDM clinic

The AHEAD study CDM clinic was located in the primary care clinic at an urban hospital. Treatment goals were to engage patients in longitudinal addiction care, to facilitate access to community resources including specialty addiction care, to communicate with caregivers including primary care providers and to re-engage patients with care after relapse or loss to clinical follow-up (Saitz et al., 2008). The CDM system components in the clinic included: (1) multidisciplinary team composed of a nurse care manager, social worker, internist with addiction specific skills, and psychiatrist; (2) informal linkages to community addiction treatment; (3) a shared electronic medical record with primary care and other medical clinicians; (4) appointment reminders and proactive callbacks; and (5) availability for drop-in care.

The CDM clinic provided some services on-site and facilitated access to needed services provided elsewhere. Addiction specific components of care available in the CDM clinic included negotiation of treatment plans, MET, a primary care adaptation of relapse prevention counseling (Friedmann, Saitz, & Samet, 1998), addiction pharmacotherapy (buprenorphine, naltrexone, and/or acamprosate), and referral to methadone maintenance treatment, specialty addiction treatment, and mutual help groups. Psychiatric assessment and treatment as well as case management for concrete needs such as food, transportation, and housing were also available. Although the CDM clinic was co-located with a primary care practice, it provided short-term medical care but not primary care. Instead the nurse care manager facilitated access to primary care, reminded patients to attend or complete evaluations for medical problems with the primary care physician, and coordinated medical and addiction treatment. Although the CDM clinic provided a somewhat diverse set of services, all were focused on improving addictions. All of these services have been conceptualized as such (Saitz et al., 2008).

The initial addiction, medical, and psychosocial clinical assessment in the CDM clinic included feedback, preventive services, initiation of addiction, short-term mental health and medical care, and additional referrals as needed. While there were overall guidelines for care, treatment was individualized based upon a patient's needs, resources (e.g., insurance), and preferences. After the care plan was in place, the nurse care manager kept in contact with the patient to assess needs and help with relapse prevention, facilitate referrals and appointments, and encourage use of CDM services.

2.3. Measures

2.3.1. CDM quality measures

Our analytic models assessed the relationship between CDM quality and addiction outcomes. Three measures of CDM quality were used—a summary of each measure is presented in Table 1. The first quality measure, engagement with CDM clinic care, was based on the Washington Circle (WC) quality measure for outpatient addiction treatment engagement (Garnick et al., 2002). Because we were interested in evaluating care specifically from the CDM clinic, the WC engagement measure was adapted to only include CDM clinic visits rather than outpatient addiction treatment visits. Engagement with the CDM clinic care was defined as two or more visits to the CDM

Table 1
Description of measures of the quality of CDM care for AOD dependence.

Measures	Definition	Analytic sample	Exposure categories
Engagement[a]	At least two visits to the CDM clinic within 30 days of initiation of care (initiation = two visits within 14 days of study enrollment)	Subjects with any follow-up data (n=553)[b]	Engagement; No engagement; No access to the CDM clinic
PACIC-CDM clinic[c]	Measure of the degree that care delivered by the CDM clinic is aligned with core components of the chronic care model[c]	Subjects assigned access to the CDM clinic with 12-month follow-up data (n=249)[d]	PACIC-CDM clinic summary score (tertiles)
PACIC-any[e]	Measure of the degree that care for addictions delivered by any healthcare provider is aligned with core components of the chronic care model	Subjects who received any care for addictions from any healthcare provider since study enrollment with 12-month follow-up data (n=451)[f]	PACIC summary score (tertiles)

[a] Based upon the WC measure of treatment engagement.
[b] Alcohol and drug addiction severity analyses limited to subjects with alcohol dependence ("alcohol subsample"; n=409) and drug dependence ("drug subsample"; n=458), respectively.
[c] Patient Assessment of Chronic Illness Care (PACIC).
[d] Alcohol subsample: n=184 and drug subsample: n=208.
[e] Any care for AOD problems includes counseling, medication, groups or other treatments provided in any healthcare setting, such as detox, hospital, emergency room, or office by any healthcare provider, including doctors, nurses, social workers, or counselors.
[f] Alcohol subsample: n=320 and drug subsample: n=378.

clinic within 30 days of "initiation" of care [initiation was at least one visit within 14 days of the first ("index") CDM clinic visit] (Kim et al, 2011). This visit frequency-based utilization measure was determined prospectively using electronic medical records adapted for the clinic. Three exposure groups for the engagement analyses were developed: (1) subjects assigned (randomly to) access to the CDM clinic who engaged, (2) those assigned access to the CDM clinic who did not engage, and (3) those (randomly) assigned to not have access to the CDM clinic. The rationale for including this third group for comparison is that the specifications of the WC measure of engagement do not include any mention of access to services; patients who receive specified services qualify as having engaged, and those who do not (regardless of the reason) are categorized as not having engaged. Consistent with that approach we did not restrict analyses to those with access to the CDM clinic.

Both the second and third measures of CDM quality were assessed using the Patient Assessment of Chronic Illness Care (PACIC), a widely used 20-item patient-completed survey that measures the extent to which care is aligned with core features of the chronic care model (Table 2) (Glasgow et al., 2005). Higher scores indicate care with more core features of the chronic care model. Specifically, the second quality measure was an assessment of the extent to which care received from the CDM clinic contained core components of the chronic care model. This measure (PACIC-CDM clinic) was assessed with a subset of questions from the PACIC survey (to minimize repetitiveness and subject burden during the interview and to focus on concepts most relevant to the services provided by the CDM clinic while ensuring that at least one question was asked about each core component of the chronic care model (patient activation, delivery system design/decision support, goal setting/tailoring items, problem-solving/contextual issues, and proactive follow-up/coordination of care). Because the questionnaire asks specifically about care provided by the CDM clinic, it was only administered to subjects assigned to have access to the CDM clinic and recalled attending it.

The third quality measure was an assessment of the extent to which care from any healthcare source was aligned with the chronic care model using the PACIC questionnaire (the complete instrument). This quality measure (PACIC-any), unlike the PACIC-CDM clinic measure, was assessed for all subjects who received any care related to AOD dependence since study enrollment. Any care included "talking or counseling, medication, groups or other treatments" provided in any healthcare setting, such as "detox, hospital, emergency room, or office" by any "healthcare provider, including doctors, nurses, social workers, counselors or others."

Both PACIC measures were determined at a 12-month interview after study entry. Given the absence of well-established PACIC score cut-offs and to avoid an assumption of linearity, both PACIC-CDM clinic and PACIC-any scores were categorized based on tertiles of the distributions. Although the PACIC was developed to assess the CDM of medical conditions, questions in the PACIC ask about highly valued components of addiction care, such as shared decision making (Institute of Medicine, 2006), proactive follow-up (McKay, 2009), "encouragement to go to a specific group" (e.g., 12-step), and help with planning "to take care of my illness even in hard times" (e.g., relapse prevention counseling).

In sum, the first two quality measures, engagement with the CDM clinic and PACIC-CDM clinic, were an assessment of care provided by the CDM clinic, whereas the third quality measure, PACIC-any, reflected care provided by any healthcare provider, AHEAD study-related or otherwise.

2.3.2. Outcome measures

Outcomes were assessed during interviews at 3, 6 and 12 months. Abstinence, the primary outcome, was defined as no opioid or stimulant use, or heavy drinking (five drinks or more per day for men; four or more for drinks per day for women) in the past 30 days. Opioid and stimulant use were assessed by the substance use questions of the Addiction Severity Index (ASI) (McLellan et al., 1992) and heavy drinking by the 30-day timeline follow back (Sobell & Sobell, 1995) calendar method. Secondary outcomes included alcohol and drug addiction severity, assessed by the alcohol and drug composite scores of the ASI. Three past 3-month addiction treatment variables (dichotomous) were examined as potential intermediate outcomes of the receipt of quality CDM: (1) specialty addiction treatment defined as outpatient or inpatient addiction treatment excluding detoxification; (2) addiction pharmacotherapy defined as medication to prevent drinking or drug use, help cut-down, or quit, excluding medication for detoxification; and (3) any mutual help group attendance.

2.4. Analytic strategy

Descriptive statistics were used to portray the study sample at baseline. We examine the relationship between measures of CDM quality and each outcome of interest by fitting separate multivariable longitudinal regression models. Generalized estimating equation (GEE) logistic regression models were used to model binary outcomes (i.e., abstinence, utilization of any specialty addiction treatment, any addiction pharmacotherapy, and any mutual help group attendance). Because the distributions of alcohol and drug ASI scores were non-normal and a suitable transformation was not identified, each variable was categorized into multiple ordered categories. The alcohol ASI composite score cutpoints were: 0; >0 to 0.1; >0.1 to 0.2; >0.2 to 0.04; >0.4 to 0.6; and >0.6. The drug ASI composite score cutpoints were: 0; >0 to 0.1; >0.1 to 0.2; >0.2 to 0.3; and >0.3). The ordered categorical data were then analyzed using GEE proportional odds models. The

Table 2
Summary of item content for Patient Assessment of Chronic Illness Care (PACIC).[a]

Chronic care model core component	Description	PACIC survey question
Patient activation	Actions that solicit patient input and involvement in decision-making	1) Asked for my ideas when my healthcare provider made a treatment plan[b] 2) Given choices about treatment to think about[b] 3) Asked about problems with medications
Delivery system design, decision support	Actions that organize care and provide information for patients to enhance their understanding of care	4) Given written list of things to do to improve my health[b] 5) Satisfied that care was well-organized[b] 6) Shown how what I did to take care of my illness influenced my condition
Goal setting, tailoring items	Acquiring information for and setting of specific collaborative goals	7) Asked to talk about my goals in caring for my illness[b] 8) Helped to set specific goals to improve my eating or exercise 9) Given a copy of my treatment plan 10) Encouraged to go to a specific group or class to help you cope with your addiction 11) Asked questions, either directly or on a survey, about your health habits
Problem-solving, contextual issues	Considering potential barriers and the patient's social and cultural environment in making treatment plans	12) Sure that my doctor/nurse caring for my *addiction* thought about my values and my traditions when they recommended treatment[b] 13) Helped to make treatment plan that I could do in my daily life[b] 14) Helped to plan ahead so I could take care of my illness even in hard times 15) Asked how *addiction* affects my life.
Proactive follow-up, coordination of care	Arranging care that extends and reinforces office-based treatment and making proactive contact with patients to assess progress and coordinate care	16) Contacted after a visit to see how things were going 17) Encouraged to attend programs in the community that could help you 18) Referred to a dietician, health educator, or counselor 19) Told how visits with other types of doctors or health professionals helps my treatment 20) Asked how visits with other doctors or health professionals were going[b]

[a] PACIC is a 20-item instrument, response scores range from 1 to 5 with higher scores indicating care that is more consistent with the chronic care model. The summary score is the mean score of all individual survey items. The word "addiction" was used in place of "my illness".
[b] PACIC questions used to assess the quality of care from the CDM clinic based upon services provided in the CDM clinic.

odds of lower scores (i.e., better outcomes), were modeled for both cases. The GEE approach was used to account for the correlation from using repeated observations from the same subject over time. An independence working correlation was used and empirical standard errors from the GEE approach are reported for all analyses.

Different analytic samples were used based on the independent variables and outcomes analyzed. A summary of the criteria used to comprise the analytic samples is presented in Table 1. Subjects with at least one follow-up interview were eligible for all analyses. Analyses of alcohol and drug severity included only subjects with any alcohol dependence ($n=409$) and any drug dependence ($n=458$), respectively at study entry. There were additional criteria for the PACIC-CDM clinic and PACIC-any analyses. Both analyses included subjects who completed those interviews (done at 12-months). The PACIC-CDM clinic analyses, however, only included those who had been randomly assigned to have access to the CDM clinic and recalled attending it. Those without access to the CDM clinic were not included. Analyses of care provided by any health care source (PACIC-any) measure consisted of those who reported receipt of any care for AOD dependence since study enrollment with no distinction made to indicate whether subjects were in the intervention or control arms of the randomized study.

All regression models were fit with a single main independent variable of interest and the following covariates: age, sex, race/ethnicity, an indicator variable for time since enrollment (3, 6, or 12 months), homelessness (any night on the street or in a shelter, past 3 months), and moderate to severe depression [Patient Health Questionnaire (PHQ-9) score ≥ 10; Kroenke, Spitzer, & Williams, 2001]—the latter two were modeled as time-dependent variables. Alcohol and drug severity analyses also included baseline alcohol and drug severity, respectively. Prior to regression modeling, potential collinearity among covariates was assessed by calculating the correlation between all independent variables and covariates, and no pair of variables had a Spearman correlation >0.40. Due to the exploratory nature of the analyses, no adjustments were made for multiple comparisons. However, pair-wise comparisons were not made unless the global p-value for the CDM measure was statistically significant ($p<.05$). All statistical analyses were performed using SAS version 9.2 (SAS Institute, Inc., NC, USA).

3. Results

3.1. Sample characteristics

This study's analytic samples were derived from the AHEAD study randomized trial. Among the 655 eligible adults, 563 (87%) were randomized in the AHEAD intervention trial. Ninety-eight percent of those randomized completed at least one follow-up interview (89%, 87%, and 95% at 3, 6, and 12-month interviews, respectively). There was no significant difference between the proportion of the intervention and control groups interviewed at each follow-up.

Among the 563 subjects enrolled in the AHEAD intervention trial, 553 (98%) had at least one follow-up interview (analytic sample for the CDM clinic engagement analyses). Among the 270 subjects assigned access to the CDM clinic, 249 (92%) completed a 12-month follow-up interview, attended the clinic and recalled doing so comprising the analytic sample for PACIC-CDM clinic analyses. The analytic sample for the "any care" (PACIC-any) analyses was 451 (those who completed the 12-month interview [532] and reported receipt of care for AOD dependence from any healthcare source since study enrollment).

The analytic sample for the engagement analyses (Table 3) had the following characteristics: men (73%); non-white (53%); homeless (59%); had both alcohol and drug dependence (66%); a comorbid substance abuse-related medical condition (65%); and moderate to severe depression (84%). At baseline, a low percentage had past 3-month utilization of specialty treatment (34%) or addiction pharmacotherapy (6%), yet nearly half had mutual help group attendance. Approximately a quarter of the subjects with access to the CDM clinic met criteria for engagement (23%). Approximately a third of the sample was abstinent at follow-up (35%, 34%, and 43% at 3, 6, and 12-month interviews, respectively) (Table 4).

Table 3
Characteristics of subjects with AOD dependence enrolled in a trial of access to primary care based CDM care with follow-up data ($N=553$).

Baseline	n (%)[a]
Age, median years (25th, 75th quartiles)	39 (29, 46)
Male	404 (73)
Race/ethnicity , non-white	293 (53)
Homeless, any (past 3 month)	327 (59)
Dependence, type[b]	
Alcohol only	65 (12)
Drug only	123 (22)
Drug and alcohol	365 (66)
Alcohol addiction severity, mean (std)[c]	0.63 (0.25)
Drug addiction severity, mean (std)[d]	0.34 (0.12)
Heroin, any (past 30 days)	329 (60)
Cocaine, any (past 30 days)	375 (68)
Heavy alcohol, any (past 30 days)	433 (78)
Depression, moderate or severe[e]	460 (84)
Substance disorder-related medical condition[f], lifetime	358 (65)
Specialty addiction treatment, any (past 3 month)	189 (34)
Addiction pharmacotherapy, any (past 3 month)	33 (6)
Mutual help group attendance, any (past 3 month)	265 (48)
Engagement with CDM clinic care[g]	127 (23)
PACIC-CDM clinic score, tertiles[h, i]	
Highest	4.29–5
Middle	3.43–<4.29
Lowest	1–<3.43
PACIC-any score, tertiles[h, j]	
Highest	3.70–5
Middle	2.95–<3.70
Lowest	1–<2.95

[a] Unless otherwise indicated.
[b] Composite International Diagnostic Interview–Short Form.
[c] Among subjects with alcohol dependence and recent heavy alcohol use, $n=409$.
[d] Among subjects with drug dependence and recent drug use, $n=458$.
[e] PHQ-9 score ≥10.
[f] Includes seizures, heart failure, atrial fibrillation, rapid heat beat, hepatitis, cirrhosis,peripheral neuropathy, cancer of mouth/esophagus/stomach, skin infections, pneumonia, tuberculosis, gastritis, pancreatitis, anemia, septic arthritis, endocarditis,or blood clots.
[g] Among subjects randomly assigned access to the CDM clinic, $n=270$.
[h] PACIC score range: 1–5 with higher scores indicating care with more core features of the chronic care model.
[i] Measure of the degree that care delivered by the CDM clinic is aligned with core components of the chronic care model.
[j] Measure of the degree that care for addictions delivered by any healthcare provider is aligned with core components of the chronic care model.

3.2. Multivariable regression results

No significant group differences were detected for the effect of engagement in CDM on abstinence or drug and alcohol addiction severity (Table 4).

Receipt of higher quality CDM from the clinic as reflected by higher PACIC-CDM scores was significantly associated with lower drug addiction severity (global p-value=.03) [adjusted odds ratio (AOR) 1.71; 95% confidence interval (CI) 1.13, 2.59, highest vs. middle tertile]. No significant associations were detected between PACIC-CDM clinic scores and abstinence or alcohol addiction severity.

Higher PACIC-any scores, reflecting the quality of care received anywhere, were associated with higher odds of abstinence (global p-value<.001) (AOR 1.99; 95% CI: 1.34, 2.95, highest vs. lowest tertile; AOR 1.75; 95% CI: 1.24, 2.48, highest vs. middle tertile) and lower alcohol severity (global p-value=.02) (AOR 1.68; 95% CI: 1.16, 2.45, highest vs. middle tertile). Higher PACIC-any scores appeared to be associated with lower drug addiction severity but the global association was not statistically significant ($p=.09$).

Upon examination of secondary outcomes, all measures of quality CDM care (Table 1) were significantly associated with specialty addiction treatment utilization (Table 5). Engagement was also associated with higher odds of addiction pharmacotherapy (AOR 3.55; 95% CI: 2.02, 6.25) but not with mutual help group attendance. Each PACIC measure, reflecting care from the CDM clinic (PACIC-CDM clinic) or any healthcare source (PACIC-any) was associated with higher odds of mutual help group attendance, but neither was significantly associated with utilization of addiction pharmacotherapy.

4. Discussion

To date, the most widely accepted measures of the quality of addiction care are based on visit frequency. In this sample of adults with AOD dependence, a measure of CDM care quality based on visit frequency—engagement—was not associated with abstinence or addiction severity despite being associated with receipt of specific addiction treatments (e.g., pharmacotherapies). However, a self-reported quality measure assessing alignment with the chronic care model was associated with several addiction outcomes. Specifically, those who received care from any healthcare source with core features of the chronic care model were more likely to be abstinent from heroin, cocaine, and heavy drinking and among those with alcohol dependence to have lower alcohol addiction severity. Those who received higher quality care specifically from the CDM clinic were more likely to have lower drug addiction severity (Table 6).

The lack of significant associations between CDM engagement and abstinence or addiction severity is notable. Engagement with CDM care as measured by visit frequency is likely a necessary minimum first step that can lead to receipt of known efficacious treatments; however in this cohort, adults with comorbidity and social ills (e.g., homelessness), it was not sufficient for improving addiction outcomes.

Findings were clearer and more consistent when the predictor of interest was the self-report quality measure of alignment with the chronic care model regardless of the source of care. Analyses examining care quality from any health care source are informative because individuals with AOD dependence often receive episodic acute care services (e.g., in detoxification units or emergency departments) rather than longitudinal comprehensive chronic care. Among patients with chronic medical conditions including diabetes mellitus, asthma, coronary artery disease, and chronic pain, higher PACIC scores

Table 4
Outcome variable distributions by time point, n (%).

Variable	3 months ($n=500$)	6 months ($n=487$)	12 months ($n=532$)
Abstinence	177 (35)	164 (34)	229 (43)
Alcohol addiction severity[a,b]			
0	28 (8)	41 (11)	53 (14)
>0–0.2	122 (33)	121 (33)	142 (36)
>0.2–0.4	95 (26)	103 (28)	88 (22)
>0.4–0.6	61 (18)	43 (12)	61 (16)
>0.6	62 (17)	56 (15)	49 (13)
Drug addiction severity[a,c]			
0	15 (4)	20 (5)	37 (8)
>0–0.1	125 (31)	136 (34)	167 (38)
>0.1–0.2	130 (32)	119 (30)	126 (29)
>0.2–0.3	70 (17)	76 (19)	64 (15)
>0.3	70 (17)	51 (13)	47 (11)
Specialty addiction treatment[d]	280 (56)	226 (47)	218 (41)
Addiction pharmacotherapy[d]	86 (17)	83 (17)	97 (18)
Mutual help group[d]	275 (55)	268 (55)	294 (55)

[a] Addiction Severity Index.
[b] Subjects with alcohol dependence and recent heavy use who completed interviews at 3-months ($n=369$), 6-months ($n=364$), and 12-months ($n=394$).
[c] Subjects with drug dependence and recent use who completed interviews at 3-months ($n=411$); 6-months ($n=402$); and 12-months ($n=441$).
[d] Past 3 months.

Table 5
Multivariable associations of three measures of high quality CDM care and addiction outcomes.[a]

Variable	Abstinence		Lower alcohol addiction severity		Lower drug addiction severity	
	Global *p*-value	AOR (95% CI)	Global *p*-value	AOR (95% CI)	Global *p*-value	AOR (95% CI)
Engagement with CDM clinic care[b]						
Yes	.2	0.94 (0.61, 1.43)	.8	1.08 (0.64, 1.82)	.4	0.86 (0.59, 1.26)
No		0.76 (0.56, 1.03)		0.94 (0.69, 1.27)		0.84 (0.64, 1.10)
Control		1		1		1
PACIC-CDM clinic (tertiles)[c]						
Highest	.1	1.71 (1.00, 2.94)	.2	1.22 (0.69, 2.13)	.03[d]	1.15 (0.71, 1.86)
Middle		1.19 (0.69, 2.03)		0.78 (0.46, 1.35)		0.67 (0.43, 1.06)
Lowest		1		1		1
PACIC-any (tertiles)[e]						
Highest	.0005[f]	1.99 (1.34, 2.95)	.02[g]	1.24 (0.81, 1.88)	.09	1.45 (1.04, 2.04)
Middle		1.13 (0.76, 1.68)		0.73 (0.49, 1.10)		1.25 (0.89, 1.76)
Lowest		1		1		1

[a] Result of separate multivariable longitudinal regression models predicting abstinence (GEE logistic regression for engagement) and lower alcohol and drug addiction severity (GEE proportional odds model). All models include age, sex, race/ethnicity, time, homelessness, and depression (PHQ-9). Alcohol and drug addiction severity models also include baseline alcohol and drug addiction severity, respectively.

[b] Engagement defined as at least two visits to the CDM clinic within 30 days of CDM clinic initiation. Analytic samples consist of subjects with follow up data ($n=553$); alcohol and drug addiction severity analyses limited to subjects with alcohol dependence ("alcohol subsample") ($n=409$) and drug dependence ("drug subsample") ($n=458$), respectively.

[c] PACIC-CDM clinic defined as the degree that care delivered by the CDM clinic was aligned with core features of the chronic care model. Analytic samples consist of subjects randomized to have access to AHEAD CDM care with 12-month interview data ($n=249$); alcohol subsample ($n=184$) drug subsample ($n=208$). PACIC-CDM clinic score tertiles: highest 4.29–5, middle: 3.43–<4.29, lowest: 1–<3.43.

[d] Highest vs. middle PACIC-CDM clinic tertile predicting lower drug severity.

[e] PACIC-any defined as the degree that care delivered by any healthcare source is aligned with core features of the chronic care model. Analytic samples consist of subjects with 12-month interview data who received any care for alcohol or other drug dependence since study entry (n=451); alcohol subsample (n=320); drug subsample (n=378). PACIC-any tertiles: highest 3.70–5; middle: 2.95 –<3.70, lowest: 1–<2.95.

[f] Highest vs. middle PACIC-any tertile predicting abstinence AOR 95% CI: 1.75 (1.24, 2.48). Highest vs. lowest PACIC-any tertile predicting abstinence AOR 95% CI: 1.99 (1.34, 2.95).

[g] Highest vs. middle PACIC-any tertile predicting lower alcohol severity AOR 95% CI: 1.68 (1.16, 2.45).

(indicating care with more features of the chronic care model) were associated with more self-management behaviors (e.g., regular exercise; Schilinger, Wang, Handley, & Hammer, 2009) and higher quality of life (Schmittdiel et al., 2007). Our results suggest that receipt of care for AOD dependence that includes components of the chronic care model is also important for addiction treatment effectiveness.

Although the quality of CDM care from any health care source was significantly associated with abstinence and alcohol addiction severity, analyses of drug addiction severity did not quite achieve statistical significance. We did find, however, that drug addiction severity was significantly associated with the quality of care from the CDM clinic. It is possible that this finding was related to the way that services were organized in the CDM clinic. For example, the clinic provided some psychiatric services on-site and facilitated linkage to other psychiatric services. This may have been particularly helpful for those with drug dependence because of the higher mental health comorbidity rate for patients with drug dependence (Grant et al., 2004). In addition, since drug dependence is more difficult to treat without addiction pharmacotherapy, we postulated that greater access to addiction pharmacotherapy in the CDM clinic may have contributed to these findings. However, we did not find that the quality of care from the CDM clinic or from any source of addiction treatment was associated with higher odds of addiction pharmacotherapy receipt.

These differences should not be overstated and are only hypothesis-generating. Although there were differences in these analyses, overall, our main findings are that regardless of the source of care, the quality of CDM care matters and that a quality assessment based upon the content of care appears to have predictive validity but a visit-based, frequency measure does not. The fact that the results of

Table 6
Association of receipt of high quality CDM and utilization of specialty addictiontreatment, addiction pharmacotherapy, and mutual help group.

	Specialty addiction treatment		Addiction pharmacotherapy		Mutual help12-step group	
	Global *p*-value	AOR (95% CI)	Global *p*-value	AOR (95% CI)	Global *p*-value	AOR (95% CI)
Engagement with CDM						
Yes	0.001	2.34 (1.51, 3.64)	0.0008	3.55 (2.02, 6.25)	0.2	1.18 (0.74, 1.87)
No		1.24 (0.94, 1.64)		1.50 (0.99, 2.27)		0.81 (0.59, 1.09)
Control		1		1		1
PACIC-CDM clinic (tertiles)						
Highest	0.004	2.13 (1.31, 3.45)	0.5	1.33 (0.70, 2.49)	0.02	1.88 (1.12, 3.14)
Middle		2.00 (1.21, 3.30)		0.93 (0.49, 1.76)		2.09 (1.22, 3.60)
Lowest		1		1		1
PACIC-any (tertiles)						
Highest	0.002	1.83 (1.30, 2.59)	0.4	1.37 (0.84, 2.26)	0.007	1.86 (1.26, 2.75)
Middle		1.18 (0.83, 1.67)		1.27 (0.77, 2.09)		1.49 (1.02, 2.18)
Lowest		1		1		1

Results of separate multivariable logistic regression models for each outcome and each main independent variable. Specialty addiction treatment defined as outpatient or inpatient addiction treatment excluding detoxification; addiction pharmacotherapy as medication to prevent drinking or drug use, help cut-down, or quit (not for detoxification); and mutual help, 12-step (e.g., AA) groups.

All models include age, sex, race/ethnicity and the following time-varying covariates: homeless (Y/N), depression (PHQ-9), alcohol and drug severity, time since study enrollment (3, 6, or 12 months).

Analytic sample sizes: engagement with CDM clinic care=553, PACIC-CDM clinic=249, PACIC-any=451.

utilization analyses were the same for each PACIC measure, (i.e., significant associations for specialty addiction treatment and mutual help group attendance but not addiction pharmacotherapy) supports the finding that both PACIC measures assessed similar constructs regardless of the source of care. Both PACIC measures seem to be assessing something different than the engagement measure as indicated by the engagement association with addiction pharmacotherapy but not mutual help group attendance.

All quality measures were associated with higher odds of utilization of specialty addiction treatment. This is not insignificant because successful referrals to "off-site" addiction treatment providers with whom the CDM clinic did not have formal referral relationships can be particularly challenging (Gurewich, Sirkin, & Shepard, 2012) and linkage to outside resources is an important component of chronic care models (Wagner et al., 1996).

This study adds to the literature supporting the benefits of longitudinal, integrated care for patients with AOD dependence by showing that the quality of CDM can contribute to improved outcomes. (O'Toole et al. 2011) found that primary care with elements of the chronic care model customized for homeless patients, of whom 70% had alcohol abuse, was associated with better medical outcomes (blood pressure, glycemic control, and lipid levels). Alcohol use outcomes were not reported. Among patients in addiction treatment, (Chi, Parthasarathy, Mertens, & Weisneret 2011) found that continuing care, defined as having yearly primary care and specialty addiction and psychiatric care when needed, was associated with abstinence over a 9-year follow-up period. The current study contributes to this literature by examining addiction outcomes from a type of care that may facilitate receipt of primary, psychiatric, and specialty care that was also structured to provide self-management support and other elements of the chronic care model.

This study has several limitations. Due to the study's observational design, it is possible that patients who were abstinent (or had lower addiction severity) tended to rate treatment more favorably. However, we used prospectively collected data to assess addiction outcomes over a 12-month study follow up period adjusted for potential determinants of care (i.e., gender, race/ethnicity, homelessness, and depression). In addition, the PACIC is an instrument that assesses implementation of the chronic care model rather than simply patient satisfaction. Nonetheless we cannot exclude the possibility that unmeasured factors influenced which participants rated treatment with high PACIC scores and also influenced addiction outcomes. We did not have information about the type of addiction treatment participants were rating. However, the question of interest was not the specific treatment modality but rather receipt of care with core features of the chronic care model. Given that there are few data about CDM for AOD dependence, these findings serve as a "proof of concept" study. Future study should examine types of care that deliver high quality CDM.

Another limitation of this study is the method used to model exposure and outcome. Assessment of CDM quality was done at the 12-month study interview and addiction outcomes were measured prospectively at 3, 6, and 12-month interviews. Subjects may have rated care that was received after the addiction outcome was measured. Studies are needed to prospectively examine the receipt of care delivered before the assessment of addiction outcomes. Finally, generalizability is another consideration. The majority of participants were recruited from a detoxification unit and most participants had social and psychiatric co-morbidities such as homelessness and depression. This study's findings may not generalize to primary care patients seeking addiction treatment or individuals with less social and psychiatric comorbidity. On the other hand, it is precisely these sorts of populations who need and could potentially benefit from high quality integrated and coordinated care.

This study's strengths include prospective data collection and a high proportion assessed at follow-up. In addition to process measures of care such as receipt of addiction treatment and addiction pharmacotherapy, we examined outcomes with clear clinical significance (i.e., abstinence and severity). In addition, we used a measure based on the NCQA/HEDIS quality performance measure for addiction treatment engagement to define engagement with the particular treatment of interest in this study, CDM. In addition, this was an innovative analysis using a tool (PACIC) not often applied to addictions care that may be increasingly used to evaluate the quality of care consistent with CDM in the patient-centered medical home.

Given the need for improvement of the quality of care for patients with addictions, CDM shows some promise for improving patient receipt of effective treatments, thereby improving outcomes. These data support efforts to improve care for patients with addictions that integrate medical, mental health and addiction services, provide longitudinal coordinated care and care that pro-actively follows patients. These data also suggest that as such efforts become more common, attention will need to be paid to the quality of care—how it reflects key features of the chronic care model—to assure the effectiveness of primary care based CDM for AOD dependence.

In conclusion, this study provides empirical data that suggest that receipt of quality CDM for AOD dependence is associated with improved processes of care and better outcomes. Furthermore, our results suggest that widely used visit-based frequency measures may be inadequate for capturing characteristics of CDM that are associated with better outcomes, characteristics better captured by self-report measures of exposure to elements of chronic care management.

Acknowledgments

We presented an earlier version of the manuscript as an oral presentation at the Addiction Health Services Research Conference, Fairfax, VA on October 4, 2011.

This study was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) grant R01 AA010870 and the National Institute on Drug Abuse (NIDA) grant R01 DA010019. This work was also supported in part by the Boston University Clinical Translational Science Institute from the National Center for Research Resources (UL1-RR025771). The sponsors had no further role in the study design; in the collection, analysis, and/or interpretation of the data; in the writing of the report; or in the decision to submit the paper for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDA, the NIAAA, or the National Institutes of Health.

References

American Academy of Family Physicians (AAFP), American Academy of Pediatrics (AAP), American College of Physicians (ACP), American Osteopathic Association (AOA). (2011). *Guidelines for Patient-Centered Medical Home Recognition and Accreditation Programs.* http://www.acponline.org/running_practice/pcmh/understanding/educ-joint-principles.pdf. Accessed February 2, 2011.

Blonde, L. (2000). Disease management approaches to type 2 diabetes. *Managed Care*, *9*, 18–23.

Chi, F. W., Parthasarathy, S., Mertens, J. R., & Weisner, C. M. (2011). Continuing care and long-term substance use outcomes in managed care: Early evidence for a primary care-based model. *Psychiatric Services*, *62*, 1194–1200.

Cohen, E., Feinn, R., Arias, A., & Kranzler, H. R. (2007). Alcohol treatment utilization: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug and Alcohol Dependence*, *12*, 214–221.

Croghn, T. W., & Brown, J. D. (2010). Integrating mental health treatment into the patient centered medical home. Rockville, MD: Agency of Healthcare Research and Quality Publication No. 10-0084-EF.

De Alba, I., Samet, J. H., & Saitz, R. (2004). Burden of medical illness in drug and alcohol dependent persons without primary care. *American Journal on Addictions*, *13*, 33–45.

Friedmann, P. D., Lemon, S. C., Stein, M. D., & D'Aunno, T. A. (2003). Accessibility of addiction treatment results from a national survey of outpatients substance abuse treatment organization. *Health Services Research*, *38*, 887–903.

Friedmann, P. D., Saitz, R., & Samet, J. H. (1998). Management of adults recovering from alcohol or other drug problems: Relapse prevention in primary care. *Journal of the American Medical Association*, *279*, 1227–1231.

Garnick, D. W., Lee, M. T., Chalk, M., Gastfriend, D., Horgan, C. M., McCorry, F., et al. (2002). Establishing the feasibility of performance measures for alcohol and other drugs. *Journal of Substance Abuse Treatment Services*, *23*, 375–385.

Gilbody, S., Bower, P., Fletcher, J., Richards, D., & Sutton, A. J. (2006). Collaborative care for depression: A cumulative meta-analysis and review of longer-term outcomes. *Archives of Internal Medicine*, *166*, 2314–2321.

Glasgow, R. E., Wagner, E., Schaefer, J., Mahoney, L. D., Reid, R. J., & Greene, S. M. (2005). Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). *Medical Care*, *43*, 436–444.

Grant, B., Stinson, F. S., Dawson, D. A., Chou, S. P., Dufour, M. C., Compton, W., et al. (2004). Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, *61*, 807–816.

Gurewich, D., Sirkin, J. T., & Shepard, D. S. (2012). On-site provision of substance abuse treatment services at community health center. *Journal of Substance Abuse Treatment*, *42*, 339–345.

Institute of Medicine. (2006). Improving the quality of healthcare for mental and substance-use conditions. Washington, DC: National Academies Press.

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H. U. (1998). The World Health Organization Composite International Diagnostic Interview Short Form (CIDI-SF). *International Journal of Methods in Psychiatric Research*, *7*, 171–185.

Kim, T. W., Saitz, R., Cheng, D. M., Winter, M. R., Witas, J., & Samet, J. H. (2011). Initiation and engagement in chronic disease management care for substance dependence. *Drug and Alcohol Dependence*, *115*, 80–86.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606–613.

McAlister, F. A., Lawson, F. M., Teo, K. K., & Armstrong, P. W. (2001). A systematic review of randomized trials of disease management programs in heart failure. *American Journal of Medicine*, *110*, 378–384.

McKay, J. R. (2009). Continuing care research: What we have learned and where are we going. *Journal of Substance Abuse Treatment*, *36*, 131–145.

McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., et al. (1992). The fifth edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment*, *9*, 199–213.

McLellan, A. T., Lewis, D. C., & O'Brien, C. P. (2000). Drug dependence: A chronic medical illness: Implications for treatment, insurance, and outcome evaluations. *Journal of the American Medical Association*, *284*, 1689–1695.

Mertens, J. R., Lu, Y. W., Parthasarathy, S., Moore, C., & Weisner, C. M. (2003). Medical and psychiatric conditions of alcohol and drug treatment patients in an HMO: Comparison with matched controls. *Archives of Internal Medicine*, *163*, 2511–2517.

National Committee for Quality Assurance. (2011). Patient centered medical home (PCMH) overview white paper, http://www.ncqa.org/tabid/1302/Default.aspx accessed January 31, 2012

Neumeyer-Gromen, A., Lampert, T., Stark, K., & Kallischnigg, G. (2004). Disease management programs for depression: A systematic review and meta-analysis of randomized controlled trials. *Medical Care*, *42*, 1211–1221.

O'Toole, T. P., Buckei, L., Bourgault, C., Blumen, J., Redihan, S. G., Jiang, L., et al. (2011). Applying the chronic care model to homeless veterans: Effect of a population approach to primary care on utilization and clinical outcomes. *American Journal of Public Health*, *100*, 2493–2499.

Saitz, R., Larson, M. J., LaBelle, C., Richardson, J., & Samet, J. H. (2008). The case for chronic disease management for addiction. *Journal of Addiction Medicine*, *2*, 55–65.

Samet, J. H., Friedmann, P., & Saitz, R. (2001). Benefits of linking primary medical care and substance abuse services: Patient, provider, and societal perspectiveness. *Archives of Internal Medicine*, *161*, 85–91.

Schilinger, D., Wang, F., Handley, M., & Hammer, H. (2009). Effects of self-management support on structure, process, and outcomes among vulnerable patients with diabetes. *Diabetes Care*, *32*, 559–566.

Schmittdiel, J., Mosen, D. M., Glasgow, R. E., Hibbard, J., Remmers, C., & Bellows, J. (2007). Patient-Assessment of Chronic Illness Care (PACIC) and improved patient-centered outcomes for chronic conditions. *Journal of General Internal Medicine*, *23*, 77–80.

Sobell, L., & Sobell, M. (1995). Timeline followback (TLFB). User's manual. Toronto, Canada: Addiction Research Foundation.

Substance Abuse and Mental Health Services Administration. SAMHSA-HRSA Center for Integrated Health Solutions, http://www.integration.samhsa.gov/about-us/what-is-integrated-care.Accessed January 30, 2012.

Wagner, E. H., Austin, B. T., Davis, C., Hindmarsh, M., Schaefer, J., & Bonomie, A. (2001). Improving chronic illness care: Translating evidence into action. *Health Affairs*, *20*, 64–78.

Wagner, E. H., Austin, B. T., & Von Korff, M. (1996). Organizing care for patients with chronic illness. *Milbank Quarterly*, *74*, 511–544.

Watkins, K., Pincus, H. A., Tanielian, T. L., & Klein, D. J. (2003). Using the chronic care model to improve treatment of alcohol use disorders in primary care settings. *Journal of Studies on Alcohol*, *64*, 209–218.

Co-morbidity is the norm, not the exception: chronic respiratory diseases in chronic drug users

See linked article by Palmer *et al*. on pg 377

*Theresa W Kim[1], Jeffrey H Samet[1,2]

[1] Department of Medicine, Section of General Internal Medicine, Clinical Addiction Research and Education (CARE) Unit, Boston University School of Medicine/Boston Medical Center, Boston, Massachusetts, USA
[2] Department of Community Health Sciences, Boston University School of Public Health, Boston, Massachusetts, USA

*Correspondence: Dr Theresa W Kim, MD, 801 Massachusetts Ave., 2nd Floor, Boston, MA 02118, USA
Tel: +1 (617) 414-6932 Fax: +1 (617) 414-4676
E-mail: theresa.kim@bmc.org

Individuals with chronic drug use disorders frequently have medical co-morbidities.[1,2] Mertens *et al.* documented the high prevalence of chronic medical conditions and its associated significant morbidity among insured patients in a specialty alcohol and drug treatment clinic.[3] Among alcohol or drug dependent patients on the opposite end of the socioeconomic spectrum in the USA – those without primary care who were entering residential detoxification – 45% reported having a chronic medical illness.[1] One third of a comparable group reported being in fair or poor health.[4] Thus, it is fair to say that co-morbidity, among the socioeconomic spectrum of chronic drug users in care, is the norm, not the exception.

Despite an expanding literature on medical co-morbidities among drug using populations, respiratory diseases have not been well-studied. The bulk of the literature on drug users and respiratory disease has focused on infectious complications of drug use, including bacterial pneumonia, septic pulmonary embolism, and tuberculosis (TB).[5,6] These medical conditions are acute or sub-acute (e.g. TB) episodic illnesses. Chronic, non-infectious respiratory diseases such as asthma and chronic obstructive pulmonary disease (COPD) among drug users have not been well-characterised. Studies demonstrating the effectiveness of disease management programmes for chronic medical conditions such as asthma typically exclude individuals with drug use.[7] However, a better understanding of the risks of chronic respiratory diseases would inform efforts to improve the medical assessment and treatment – both therapeutic and preventive – of patients who are also drug users. Such work would address a treatment gap for a population less likely to receive quality asthma care[7] and flu vaccination.[8]

In this issue of the *PCRJ*, Palmer *et al*. report on respiratory diseases in drug users.[9] Specifically, they performed a cross-sectional analysis of the association of chronic respiratory diseases and "drug misuse" using administrative data collected from general practices in Scotland. The authors found that drug users were more likely to have a diagnosis of chronic respiratory diseases (i.e. asthma, COPD, and "respiratory system disease") than a control group matched for age, gender, and economic status. Drug users were also more likely to be prescribed respiratory medications, primarily bronchodilators and inhaled corticosteroids. These associations did not appear to be fully attributable to the high prevalence of tobacco use, since these differences persisted with adjustment for tobacco use. This study's results are consistent with others in the literature,[1-4] and Palmer *et al*. substantively contribute to this literature by performing their analysis on a national database.[9] Additionally, the use of a control group matched on key socio-demographic variables strengthens the research design.

Nonetheless, interpretation of study findings requires consideration of study limitations. First, substance type was unknown in nearly three-quarters of cases (72%), making it difficult to assess what accounts for the relationship between drug use and respiratory disease. Second, chronic respiratory disease was defined as ever having a diagnosis of asthma, COPD, or "respiratory system disease" since birth; therefore it is unknown whether the respiratory disease predated the onset of drug use. Third, although the exact number of opioid users in the study is unknown, some portion of the drug "misuser" sample consisted of patients prescribed methadone for opioid dependence from general practices, which may have required frequent office visits; therefore, respiratory diseases may have been more likely to be detected given the increased surveillance among these patients.

Despite these caveats, the study raises useful questions and suggests pragmatic implications. As noted among other studies of drug users, tobacco use in this study's sample[9] was almost universal (90%). Although the findings may not have been fully attributable to tobacco use, tobacco is still the major driver responsible for respiratory disease in this population. From a clinical perspective, given that COPD is typically under-diagnosed and under-treated in primary care,[10] these findings, although not definitive, should raise awareness that individuals with drug addiction are at high risk of chronic respiratory diseases. From a research perspective, determining whether or not contributors other than tobacco exist would be of value, as it might enable medical teams pro-actively to direct attention to other useful evaluation and treatment.

Evidence of an increased risk of chronic respiratory diseases among drug users also enhances the case for the coordination of care between addiction treatment providers and those addressing these patients' chronic medical conditions.[11] Drug treatment is an opportunity for the patient not only to address drug use but also to consider the wider health consequences of drug use including respiratory illness and linking patients to medical care.[12,13] Raising the issue with the patient and making a plan to address this common diagnosis might serve not only to improve their health and quality of life but could build the rationale for the patient to make recovery a priority. Weisner *et al*. demonstrated improved addiction outcomes

with integrated addiction and medical treatment for those with chronic medical conditions.[14]

This examination of respiratory diseases among drug users[9] adds further strength to the premise that, given the frequency of overlapping drug use and medical co-morbidities, improvements in these health domains will likely require health professionals to be cognisant of both, in the quest to address either optimally.

Conflicts of interest The authors declare that they have no conflicts of interest in relation to this article.

Contributorship Both authors have contributed significantly to the writing and editing of the manuscript, and have approved this manuscript.

Commissioned article; not externally peer-reviewed; accepted 28th October 2012; online 8th November 2012

http://dx.doi.org/10.4104/pcrj.2012.00094
Prim Care Respir J 2012; **21**(4): 364-365

References

1. De Alba, Samet JH, Saitz R. Burden of medical illness in drug- and alcohol-dependent persons without primary care. *Am J Addict* 2004;**13**(1):33-45. http://dx.doi.org/II10.1080/10550490490265307
2. Fareed A, Musselman D, Byrd-Sellers J, *et al*. On-site Basic Health Screening and Brief Health Counseling of Chronic Medical Conditions for Veterans in Methadone Maintenance Treatment. *J Addict Med* 2010;**4**(3):160-6. http://dx.doi.org/10.1097/ADM.0b013e3181b6f4e5
3. Mertens JR, Lu YW, Parthasarathy S, Moore C, Weisner CM. Medical and psychiatric conditions of alcohol and drug treatment patients in an HMO: comparison with matched controls. *Arch Intern Med* 2003;**163**(20):2511-17. http://dx.doi.org/10.1001/archinte.163.20.2511
4. Reif S, Larson MJ, Cheng DM, Allensworth-Davies D, Samet J, Saitz R. Chronic disease and recent addiction treatment utilization among alcohol and drug dependent adults. *Subst Abuse Treat Prev Policy* 2011;**6**:28. http://dx.doi.org/10.1186/1747-597X-6-28
5. O'Donnell AE, Pappas LS. Pulmonary complications of intravenous drug abuse. Experience at an inner-city hospital. *Chest* 1988;**94**(2):251-3. http://dx.doi.org/10.1378/chest.94.2.251
6. Gordon RJ, Lowy FD. Bacterial infections in drug users. *N Engl J Med* 2005;**353**(18):1945-54. http://dx.doi.org/10.1056/NEJMra042823
7. Baxter JD, Samnaliev M, Clark RE. The quality of asthma care among adults with substance-related disorders and adults with mental illness. *Psychiatr Serv* 2009;**60**(1):43-9. http://dx.doi.org/10.1176/appi.ps.60.1.43
8. Lasser KE, Kim TW, Alford DP, Cabral H, Saitz R, Samet JH. Is unhealthy substance use associated with failure to receive cancer screening and flu vaccination? A retrospective cross-sectional study. *BMJ Open* 2011;**1**(1):e000046. http://dx.doi.org/10.1136/bmjopen-2010-000046
9. Palmer F, Jaffray M, Moffat MA, Matheson C, McLernon DJ, Coutts A, Haughney J. Prevalence of common chronic respiratory diseases in drug misusers: a cohort study. *Prim Care Respir J* 2012;**21**(4):377-83. http://dx.doi.org/10.4104/pcrj.2012.00069
10. Jones R, Østrem A. Optimising pharmacological maintenance treatment for COPD in primary care. *Prim Care Respir J* 2011;**20**(1):33-45. http://dx.doi.org/10.4104/pcrj.2010.00069
11. Institute of Medicine. Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series. 1st ed. National Academies Press; 2006.
12. Samet JH, Friedmann P, Saitz R. Benefits of linking primary medical care and substance abuse services: patient, provider, and societal perspectives. *Arch Intern Med* 2001;**161**(1):85-91. http://dx.doi.org/10-1001/pubs.Arch Intern Med.-ISSN-0003-9926-161-1-ioi00065.
13. Samet JH, Larson MJ, Horton NJ, Doyle K, Winter M, Saitz R. Linking alcohol- and drug-dependent adults to primary medical care: a randomized controlled trial of a multi-disciplinary health intervention in a detoxification unit. *Addiction* 2003;**98**(4):509-16. http://dx.doi.org/10.1046/j.1360-0443.2003.00328.x
14. Weisner C, Mertens J, Parthasarathy S, Moore C, Lu Y. Integrating primary medical care with addiction treatment: a randomized controlled trial. *JAMA* 2001;**286**(14):1715-23. http://dx.doi.org/10.1001/jama.286.14.1715

Adolescents and anaphylaxis

See linked article by Gallagher *et al*. on pg 392

*Richard F Lockey[1]

[1] Distinguished University Health Professor; Professor of Medicine, Pediatrics & Public Health; Director, Division of Allergy & Immunology; Joy McCann Culverhouse Professor of Allergy & Immunology, University of South Florida College of Medicine, USA

*Correspondence: Professor Richard F Lockey, Culverhouse Professor of Allergy & Immunology, University of South Florida College of Medicine, c/o VA Hospital, 13000 Bruce B.Downs Blvd (111D), Tampa, FL 33612, USA
Tel: +1 (813) 972-7631 Fax: +1 (813) 910-4041
E-mail: rlockey@health.usf.edu

The article by Gallagher *et al*.[1] in this issue of the *PCRJ* is of particular interest, not least because the authors studied 26 adolescents (aged 13 to 19 years) at risk for anaphylaxis and did *not* find the "rebel without a cause" mentality as the dominant reason why this age group is three to four times more likely to die than children because they take risks and have accidents.[2] The movie, "Rebel Without a Cause", starring James Dean, was released by Warner Bros on 27th October 1955, less than one month after the famous actor's fatal car crash. In this controversial movie of the time, Jim Stark (played by James Dean) is a rebellious 17-year-old teenager caught up in family discord who disobeys his parents, defies the local schoolboys, and confronts the differences and conflicts between generations.[3] One would think that teenagers susceptible to anaphylaxis would possess the James Dean-like rebellious personality, thereby accounting for the fact that adolescents with this disease are at increased risk of fatal outcomes.

However, just the opposite was found in this study. The majority of episodes of anaphylaxis, primarily from food allergies (in particular, peanuts and tree nuts) but also due to fish, shellfish, sesame seed, dairy, lentils, milk, egg, tomato and other fleshy fruits, legumes and insects, horse and idiopathic, "…did not result from stereotypically irresponsible behavior (such as alcohol misuse or deliberate exposure to known allergens)."[1] The authors found that most adolescents, as did their parents, took an active role in managing their risk of anaphylaxis. Some parents and adolescents emphasised the idea that

Depression and Anxiety Diagnoses Are Not Associated with Delayed Resolution of Abnormal Mammograms and Pap Tests Among Vulnerable Women

Andrea C. Kronman, MD, MSc[1], Karen M. Freund, MD, MPH[1], Tim Heeren, PhD[2], Kristine A. Beaver, MPH[1], Mary Flynn, BA, MS[1], and Tracy A. Battaglia, MD, MPH[1]

[1]Women's Health Unit, Section of General Internal Medicine, Department of Medicine, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, MA, USA; [2]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.

BACKGROUND: Delays in care after abnormal cancer screening contribute to disparities in cancer outcomes. Women with psychiatric disorders are less likely to receive cancer screening and may also have delays in diagnostic resolution after an abnormal screening test.

OBJECTIVE: To determine if depression and anxiety are associated with delays in resolution after abnormal mammograms and Pap tests in a vulnerable population of urban women.

DESIGN: We conducted retrospective chart reviews of electronic medical records to identify women who had a diagnosis of depression or anxiety in the year prior to the abnormal mammogram or Pap test. We used time-to-event analysis to analyze the outcome of time to resolution after abnormal cancer screening, and Cox proportional hazards regression modeling to control for confounding.

PARTICIPANTS: Women receiving care in six Boston-area community health centers 2004–2005: 523 with abnormal mammograms, 474 with abnormal Pap tests.

RESULTS: Of the women with abnormal mammogram and pap tests, 19% and 16%, respectively, had co-morbid depression. There was no difference in time to diagnostic resolution between depressed and not-depressed women for those with abnormal mammograms (aHR=0.9, 95 CI 0.7,1.1) or Pap tests (aHR=0.9, 95 CI 0.7,1.3).

CONCLUSIONS: An active diagnosis of depression and/or anxiety in the year prior to an abnormal mammogram or Pap test was not associated with a prolonged time to diagnostic resolution. Our findings imply that documented mood disorders do not identify an additional barrier to resolution after abnormal cancer screening in a vulnerable population of women.

Received December 2, 2010
Revised July 5, 2011
Accepted October 10, 2011
Published online November 15, 2011

KEY WORDS: depression; cancer screening; women's health; minority populations.
J Gen Intern Med 27(4):452–7
DOI: 10.1007/s11606-011-1920-5

BACKGROUND

Depression is associated with high overall mortality, [1] including increased breast cancer mortality [2]. The underlying mechanism explaining the relationship between depression and increased breast cancer mortality is unknown. A possible explanation is underuse of screening mammography, given that screening mammography reduces mortality, [3,4] and women with psychiatric disorders are less likely to receive standard preventive care,[5,6] including screening mammography.[7] Another possible explanation is delayed follow-up after an abnormality is detected on a screening test, since delayed diagnosis can reduce survival if cancer is diagnosed at a later stage. Vulnerable populations of women, as defined by low income or with racial/ethnic minority status, are less likely to receive standard preventive healthcare and therefore experience worse breast and cervical cancer outcomes. Previous studies showed less than 25% of these vulnerable women received adequate follow-up care after an abnormal cancer screening test.[8,9] In addition, the prevalence of depression is higher in vulnerable populations,[10–16] and twofold higher in women than men. National annual prevalence estimates for women range between 4 – 14%.[17,18]

Our previous work examined factors associated with delays in cancer care among a diverse population of women who received screening at urban community health centers (CHC).[19,20] We found the individual CHC site was a stronger predictor of timely resolution than race, ethnic group, language, insurance status, or age.[19,20] We then sought to examine additional patient characteristics that may identify a group of women who are particularly vulnerable to delayed resolution of abnormal cancer screening tests and might benefit from tailored interventions such as patient navigation. We hypothesized that depression would contribute to delays in

diagnostic resolution. In addition, since previous studies demonstrated that anxiety may result in more timely care,[21,22] we hypothesized that anxiety may decrease time to resolution, and modify an association between depression and time to diagnostic resolution. Thus, our objective is to determine if depression and anxiety are independent predictors of delays in care after abnormal mammograms and Pap tests in a vulnerable population of urban women.

METHODS

Study Design

This is a secondary analysis of data collected at baseline of the Boston Patient Navigation Research Program (PNRP), before a patient navigation intervention was implemented.[19] Boston PNRP partnered with six CHCs to collect retrospective medical record data for women who had abnormal mammograms or Pap tests. For this study, data were collected via chart review for psychiatric diagnoses, symptoms, and treatment for the 1-year period preceding the date of the abnormal cancer screening test.

Study Population

Eligible subjects in the baseline PNRP included adult women with an abnormal screening Pap test or mammogram performed between January 1, 2004 and December 30, 2005 at the 6 CHCs.[19] Eligible mammograms included Breast Imaging Reporting and Data System (BIRADS) scores indicating need for follow-up (BIRADS 0, 3, 4, and 5). Eligible Pap tests included cellular abnormalities indicating need for follow-up: atypical squamous cells of undetermined significance positive for human papillomavirus (ASCUS/HPV+), low-grade squamous intraepithelial lesion (LGSIL), and high-grade squamous intraepithelial lesion (HGSIL). All subjects with high-grade abnormalities were included (BIRADS 4, 5; HGSIL), and a random sample of low-grade abnormalities (BIRADS 0, 3; ASCUS/HPV+; LGSIL) to obtain approximately 100 women per site. At sites with fewer than 100 eligible cases, all eligible subjects were included. In order to prevent clustering of cases by CHC, a representative sample of equivalent numbers were selected from each center.

Data Sources

We previously described the details of how the data were retrieved from the EMR and registration databases of the six CHCs.[19] Socio-demographic and eligibility criteria were retrieved automatically, while clinical outcomes and psychiatric diagnosis and treatment data were abstracted manually by trained research staff.

Independent Variables of Interest

We categorized a subject as actively "depressed" if either of the following criteria was documented in the EMR during the 12 months preceding the abnormal test: 1) a new diagnosis of depression [ICD-9 diagnoses 311.x, 309.28, 309.0-309.1, 300.4] entered into the active problem list[1], or 2) depression diagnosis documented in the free-text "Assessment/Plan" section of office notes. By categorizing depression in this way, we ensured that depression was an active problem in the year prior to the abnormal cancer screening test. In order to further validate our definition of depression diagnosis, we abstracted depressive symptom data (depressed mood, sleep disorders, anxious feelings, fatigue, impaired concentration, appetite change, psychomotor change, social isolation, decreased interest, suicidality, hopelessness, other) from the EMR if documented as free-text in the "History of Present Illness" or "Review Of Systems" sections of provider notes, or listed in the Problem List during the specified time period. We categorized a subject as "anxious" if anxiety was documented either in the active problem list [ICD-9 diagnoses 300.0x, 300.2x] or in the free-text sections of office notes during the specified time period.

Covariate and Intervening Variables

We obtained anti-depressant medication data from the medication list in the EMR, and evidence that the patient received psychotherapy through the following indicia: presence of an office note from a therapist or psychiatrist, mention of therapy in the free text Assessment/Plan section, or referral to behavioral health from the Orders section. We then categorized subjects into one of four mutually exclusive categories based on depression treatment: 1) anti-depressants only, 2) psychotherapy only, 3) both anti-depressants and psychotherapy, or 4) no treatment. Other psychiatric diagnoses were abstracted from the Problem Lists in the EMR, including bipolar, psychoses, and substance abuse. We collapsed the race/ethnicity EMR data into four mutually exclusive categories: White, Black/African American, Hispanic/Latina, or Asian/Other. Because the age distribution is very different for women obtaining breast and cervical cancer screening (older women in the breast cancer screening group), we categorized different age categories for the two screening populations and subdivided them into three age groups that had clinical significance. We categorized primary language as English, Spanish or other. We created the following three mutually exclusive categories from the primary and secondary insurance data in the EMR: no health insurance, publicly financed health insurance (Medicare and/or Medicaid), or private health insurance.

Outcome Variables

Our primary outcome of interest was time (number of days) from index screening abnormality to diagnostic resolution. Subjects were followed for a maximum of one year. We defined d*iagnostic resolution* as either a definitive tissue diagnosis (biopsy with pathology report) or clinical evaluation (such as colposcopy) indicating no further need for evaluation. For subjects with LGSIL pathology, we considered the initial colposcopy with biopsy to be definitive, although surveillance is recommended. Similarly, for subjects with a recommended two-year surveillance for BIRADS three results, we considered diagnostic resolution as the next six-month mammogram, subtracting the six-month period for timely resolution in the

analysis[23]. Due to the long times to resolution and most of the cases resolving within 6 months, we censored this outcome at a maximum of 365 days.

Data Analysis

Women with abnormal mammograms were analyzed separately from women with abnormal Pap tests, because of marked differences between the groups in age, racial/ethnic distribution and characteristics of cancer screening test. Chi-square tests were used to compare the characteristics of depressed and not-depressed women and to compare the prevalence of depression across demographic subgroups. Next, we examined depressive symptom data in the depressed and not depressed groups to validate our definition of depression diagnosis. In the time-to-event analysis, we first compared time to resolution within 365 days between depressed and not-depressed women via the log-rank test. We then used Cox-proportional hazards modeling to predict timely resolution incorporating all of the covariates, and an interaction term for anxiety and depression. In the final models, we included only those categorical variables with a significant univariate Cox model p-value (< 0.05), excluding the interaction term because it was not significant. In secondary analyses, we modeled separate effects for depressed women with and without treatment, because treatment for depression may modify the effect of depression diagnosis. To determine if depression was associated with outcome for any subgroups of our population, we stratified our analyses by insurance, age and race/ethnicity. For women with abnormal Pap tests, the cumulative incidence curves with and without depression cross at about 120 days; therefore we modeled separate effects of depression for resolution before and after 120 days, using time-dependent indicators for depression in the Cox model. Analyses were conducted using SAS 9.1.[24] A two-sided p-value <0.05 was considered statistically significant for reporting associations.

RESULTS

Among 997 women, 523 had abnormal mammograms and 474 had abnormal Pap tests. Overall, 17% had depression and 8.5% had anxiety. Of the women with abnormal mammogram and Pap tests, 19% and 16%, respectively, had co-morbid depression. As expected, women with abnormal mammograms were older (96%>40 years old) compared to the women with abnormal Pap tests (70% between ages 18 – 30 years old). For both groups, depressed women were more likely to be on public insurance, and for women with abnormal mammograms, depressed women were less likely to be black. (Table 1) Depressed women in both study groups were more likely to have anxiety.

Depressive symptoms were statistically significantly more common among women defined as depressed compared to those who were not. The most common documented symptoms in the depressed group in decreasing frequency were: depressed mood (57% for both women with abnormal mammograms and Pap tests), sleep problems (41% and 38%), anxious feelings (33% and 15%), fatigue (32% and 18%), impaired concentration (21% and 14%), and changes in appetite (14% for either abnormal test). For depressed women with abnormal mammograms, 69% had some form of documented treatment (66% prescribed antidepressants), while 51% of depressed women with abnormal Pap tests had some treatment (45% prescribed antidepressants).

The median time to resolution was 27 days for women with abnormal mammograms, and 85 days for women with abnormal Pap tests. Figure 1 compares the time-to-diagnostic-resolution event analysis of women with and without depression diagnosis. In both of these screening groups, we note delays in diagnostic resolution, with less than half receiving a definitive diagnosis within 30 days after abnormal mammograms, and within 90 days after abnormal Pap tests. There was no difference in time to diagnostic resolution between those with and without depression diagnosis either for women with abnormal mammograms (Fig. 1a, p=0.22 from the log-rank test) or for women with abnormal Pap tests (Fig. 1b, p=0.53).

Tables 2a and b present the bivariate and multivariable findings from the Cox proportional hazard models predicting time to resolution for mammogram and Pap test abnormalities, respectively. In these analyses, a hazard ratio less than 1.0 indicates longer time to diagnostic resolution. When examining the whole study population, neither depression diagnosis nor anxiety was significantly associated with delayed diagnostic resolution for women with abnormal mammograms or abnormal Pap tests. Secondary analyses found no significant interaction between depression and anxiety for women with abnormal mammograms (p=0.64) or with abnormal Pap tests (p=0.16). The effects of depression remained non-significant when allowing for different effects in depressed women receiving treatment (aHR 0.8 (0.6, 1.1) for mammograms and aHR 1.0 (0.7, 1.4) for Pap tests), or not receiving treatment (aHR 1.0 (0.7, 1.5) for mammograms and aHR 0.8 (0.6 , 1.2) for Pap tests) for depression. Among women with abnormal Pap tests, there was no suggestion of a depression effect on time to resolution over the first 120 days of follow-up (aHR 1.0 (0.7, 1.4)); however, there was a non-significant trend toward slower resolution for depressed women after 120 days (aHR 0.6 (0.3, 1.1), p=0.096). We performed additional secondary stratified analyses in order to examine associations within subgroups of patient sociodemographics. Separate analyses by insurance status showed no depression effect for those with public or private insurance, but demonstrated a trend of delayed diagnostic resolution for those with no insurance (HR=0.3 (0.1 , 1.0) p=0.06 for mammograms and HR=0.5 (0.2 , 1.2) p=0.11 for Pap tests). There were no differences by race, ethnicity, or age.

DISCUSSION

To our knowledge, this is the first examination of the relationship between pre-existing depression with and without anxiety and time to diagnostic resolution of abnormal mammograms and Pap tests. In this vulnerable population seeking care at urban federally qualified community health centers, we found delays in diagnostic resolution after abnormal cancer screening. However, those with a depression diagnosis did not have increased delays compared to not-depressed women. In addition, there was no significant interaction between anxiety and depression.

Table 1. Characteristics of Depressed and Not Depressed With Abnormal Mammography and Pap Test Screening

	Abnormal Mammograms, N = 512				Abnormal Pap Tests, N = 468		
	Depressed (%)	Not Depressed (%)	P		Depressed (%)	Not Depressed (%)	P
N(%)	97 (19)	415 (81)			74 (16)	394 (84)	
Race/Ethnicity			0.009				0.648
White	41 (42)	149 (36)			22(30)	98 (25)	
Black	22 (23)	148 (37)			21 (28)	139(35)	
Hispanic	27 (28)	69 (17)			22 (30)	106 (27)	
Other	7 (7)	49 (12)			9 (12)	51 (13)	
Age			0.550				0.246
18-40	4 (4)	18 (4)			5 (7)	43 (11)	
40-65	84 (87)	343 (83)			37 (50)	231 (59)	
Over 65		54 (13)			32 (43)	120 (30)	
Lanuage			0.194				0.740
English	58 (60)	269 (65)			51 (69)	263 (67)	
Spanish	19 (20)	52 (13)			12 (16)	58 (15)	
Other	20 (21)	94 (23)			11 (15)	73 (19)	
Insurance			0.022				0.002
None	13 (13)	98 (24)			12 (16)	119 (30)	
Public	36 (37)	120 (29)			32 (43)	103 (26)	
Private	48 (49)	197 (47)			30 (41)	172 (44)	
Screening Abnormality							
BIRADS* Results				Pap results			0.798
0	68 (70)	278 (67)		HGSIL†	10 (14)	49 (12)	
3	26 (27)	101 (24)		LGSIL‡	64 (86)	345 (88)	
4/5	3 (3)	36 (9)					
Community Health Center							
A	23 (24)	72 (17)	0.794		16 (22)	83 (21)	0.172
B	18 (19)	86 (21)			7 (9)	42 (11)	
C	19 (20)	87 (21)			9 (12)	72 (18)	
D	11 (11)	58 (14)			4 (5)	45 (11)	
E	7 (7)	29 (7)			15 (20)	76 (19)	
F	19 (20)	83 (20)			23 (31)	76 (19)	
Other Psychiatric Diagnoses							
Anxiety	23 (24)	14 (3)	0.001		22 (30)	26 (7)	0.001
Psychosis	1 (1)	3 (1)	0.839		0 (0)	0 (0)	0.999
Bipolar	4(4)	4 (1)	0.046		1 (1)	2(<1)	0.934
Substance abuse	0 (0)	1 (<1)	0.811		4 (5)	7 (2)	0.079

*BIRADS = Breast Imaging Reporting and Data System
†HGSIL = High Grade Squamous Intraepithelial Lesion
‡LGSIL = Low Grade Squamous Intraepithelial Lesion

Our findings suggest that for women with the lengthiest delays in diagnostic resolution (> 120 days), depression may contribute to diagnostic delays of abnormal Pap tests, but not abnormal mammograms. This may reflect differences in the populations themselves, as the populations are remarkably different in age, or differences in the perceived implications of delayed diagnostic resolution. For example, a delay of several months for an abnormal mammogram may be clinically more significant than several months of delayed resolution for abnormal Pap tests. However, our results contrast from studies of cancer screening tests reporting depression being associated with less screening mammography,[2,25] but not with fewer Pap tests.

Prior research found conflicting effects between the relationship of cancer screening behavior and depression. This may be due to depression frequently being confounded with anxiety, and the potential for anxiety having opposite effects on health outcomes. For example, a meta-analysis found that depressed patients were three times less likely to be adherent to treatment recommendations for a variety of illnesses, including breast cancer, but anxiety had little effect[6]. Patten and colleagues surveyed a large population, finding that a major depressive episode was not associated with receiving a screening mammogram in the subsequent year; however, co-existing anxiety was not accounted for in that analysis.[26] Additional analytical complexities are introduced by differential reporting and documentation of depressive and anxious symptoms. Frequently, depression and anxiety symptoms are under-documented.[27] Even if symptoms are documented, there remains an analytic challenge of determining if these symptoms reflect a stable psychological "trait" versus a temporary "state," and determining which is more important in determining future health behavior and outcomes. This distinction is of particular importance when studying the outcomes of follow-up care after abnormal cancer screening, as in this study, for the negative psychological states associated with receipt of an abnormal test could adversely affect the outcome of follow-up care.

In addition to the analytical difficulties, the relationship of depression and cancer screening behavior could vary by populations of women. For example, in our study population of ethnically diverse, economically disadvantaged women, who already have other socioeconomic barriers to reaching timely resolution after abnormal cancer screening,[28,29] our findings suggest that depression does not contribute to delays in diagnostic care, except perhaps for those who both are depressed and lack insurance. It is possible that depression does delay diagnostic resolution in women not socio-

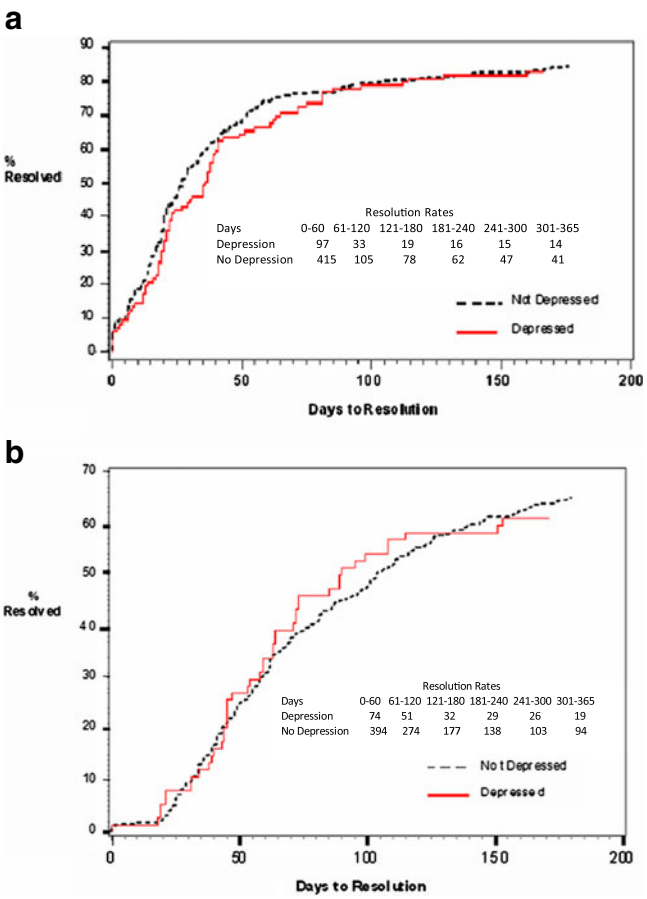


Figure 1. Time to event analysis comparison of depressed and not-depressed patients (a) Abnormal mammograms (b) Abnormal pap tests.

economically disadvantaged. Conversely, the women in our study were already engaged in the health care system, as having an abnormal cancer screening test was a mandatory inclusion criterion. If depression is a barrier to receiving screening mammography and/or Pap tests in a sub-population of women, they would not have been included in our study.

This study is limited by the relatively small sample size, limiting our power to detect statistical differences between the depressed and not-depressed groups. Using retrospective chart review to classify patients as depressed has the potential for misclassification bias in two ways, each biasing the results toward the null. First, it is possible that patients classified as "not-depressed" were truly depressed. In addition to known under-diagnosis of depression in primary care,[30] which may be differential in vulnerable populations[31] and those without insurance,[32] there may be under-documentation of it as well, since primary care providers are frequently not reimbursed if depression is the primary ICD-9 diagnosis. The accuracy of using administrative data or the EMR to determine depression diagnosis ranges from 58 – 83% depending on the contents of the EMR.[33,34] However, our 12 month depression prevalence was slightly higher than that previously reported for depression using survey tools, suggesting the possible extent to which under-diagnosis and under-reporting of depression were limited in our study.[15,35,36] Second, patients adequately treated for depression may be classified as "depressed" even though they are no longer experiencing symptoms. However, since almost all of our depressed patients had corresponding documented depressive symptoms, it is unlikely that a patient classified as depressed did not have depressive symptoms during the year preceding the abnormal cancer screening test. Although depression diagnosis documented in administrative data is more likely to correspond to severe and recurrent depression in patients,[37] we were not able to account for the clinical severity of depression or whether it was adequately treated. Finally, as our data was collected via retroactive chart review, there may be some residual confounding that we were not able to account for in our analyses.

This is the first examination of the relationship between pre-existing depression and time to diagnostic resolution of abnormal mammograms and Pap tests.[9] Previous studies focused on *screening* mammography and Pap tests. Our findings did not support the notion that pre-existing mood disorders documented in the EMR delays the time to resolution of abnormal cancer screening tests for vulnerable populations, although we did find a trend of delayed care in depressed women who were uninsured or took longer to resolve their Pap abnormalities. It is possible that the timing of depressive symptoms in relation to the time of being informed of abnormal cancer screening is a more informative predictor. Future studies should validate our results by using survey data to classify patients' depressive symptoms, their severity, and whether they reflect a stable psychological "trait" or temporary "state."

Our current findings are consistent with our previous work, in that individual characteristics of patients may be less important than system characteristics in determining the time to diagnostic resolution of abnormal mammograms or Pap tests in vulnerable women[19,20]. However, identifying which patients are most vulnerable may allow for more efficient allocation of resources, as patient-centered-medical-home models emerge and incorporate more elements of case management. Our findings imply that pre-screening the EMR for mood disorders may not be the most reliable approach to identify a group of patients at higher risk of delayed diagnostic resolution of abnormal cancer screening tests in a vulnerable population.

Table 2. Cox Proportional Hazards Analysis* Predicting Time to Resolution

	Univariate	Bivariate	Multivariate†	P
	HR (95% CI)	HR (95 % CI)	HR (95 % CI)	
Abnormal Mammograms N=538				
Depression	0.9 (0.7, 1.2)	0.9 (0.7, 1.2)	0.9 (0.7, 1.1)	0.31
Anxiety	0.9 (0.6, 1.2)	0.9 (0.6, 1.3)	1.0 (0.7, 1.4)	0.85
Abnormal Pap Tests N=415				
Depression	0.9 (0.7, 1.3)	0.96 (0.7, 1.3)	0.9 (0.7, 1.3)	0.73
Anxiety	0.9 (0.6, 1.3)	0.9 (0.6, 1.3)	1.1 (0.5, 1.1)	0.21

**Larger hazards ratios are associated with shorter time to resolution. p value is testing the HR from the multivariable model*

†Adjusted for individual community health center and degree of mammogram abnormality

‡Adjusted for individual community health center, race, language and insurance status. The cumulative incidence curves for women with and without depression cross at 120 days; therefore we modeled separate effects of depression for resolution within 120 days and after 120 days, using time-dependent indicators for depression in the Cox model

Acknowledgements: *This work was presented at the Society for General Internal Medicine National Meeting, in Miami Beach, FL, on 14 May 2009.The authors thank Ignacio De La Cruz, John Pagliaro, and Cynthia Schoettler for their support with manuscript preparation. Dr. Kronman is supported by a Cancer Control Career Development Award for Primary Care Physicians from the American Cancer Society (CCCDA-09-217-01 Kronman). This work is also supported by the National Cancer Institute (CA116892 Freund).*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Andrea C. Kronman, MD, MSc; Women's Health Unit, Section of General Internal Medicine, Department of Medicine, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, 801 Massachusetts Avenue, Suite 470, Boston, MA 02118, USA (e-mail: andrea.kronman@bmc.org).*

REFERENCES

1. **Charbonneau A, Rosen AK, Owen RR, et al.** Monitoring depression care: in search of an accurate quality indicator. Medical Care. 2004;42:522–531.
2. **Pirraglia PASP, Singer DE, Ferris TG.** Depressive symptom burden as a barrier to screening for breast and cervical cancers. J Womens Health. 2004;13:731–738.
3. **Jørgensen KJ.** Mammography reduces breast cancer mortality in women aged 39–69 years; but harms may outweigh benefits in women under 50. Evid. Based Med. 2010;15:62–63.
4. **Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L.** Screening for Breast Cancer: An Update for the U.S. Preventive Services Task Force. Ann. Intern. Med. 2009;151:727–737.
5. **Lasser KE, Zeytinoglu H, Miller E, Becker AE, Hermann RC, Bor DH.** Do women who screen positive for mental disorders in primary care have lower mammography rates? Gen. Hosp. Psychiatry. 2003;25:214–216.
6. **DiMatteo MR, Lepper HS, Croghan TW.** Depression is a risk factor for noncompliance with medical treatment: Meta-analysis of the effects of anxiety and depression on patient adherence. Arch. Intern. Med. 2000;160:2101–2107.
7. **Pirraglia PA, Sanyal P, Singer DE, Ferris TG.** Depressive symptom burden as a barrier to screening for breast and cervical cancers. J. Womens Health. 2004;13:731–738.
8. **Yabroff KR, Washington KS, Leader A, Neilson E, Mandelblatt J.** Is the Promise of Cancer-Screening Programs Being Compromised? Quality of Follow-Up Care after Abnormal Screening Results. Med. Care Res. Rev. 2003;60:294–331.
9. **Bastani RYK, Myers RE, Glenn B.** Interventions to improve follow-up of abnormal findings in cancer screening. Cancer. 2004;101:1188–1200.
10. **Breslau J, Kendler KS, Su M, Gaxiola-Aguilar S, Kessler RC.** Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. Psychol. Med. 2005;35:317–27.
11. **Cutrona C, Russell DW, Brown PA, Clark LA, Hessling RM, Gardner KA.** Neighborhood context, personality, and stressful life events as predictors of depression among African American women. J. Abnorm. Psychol. 2005;114:3–15.
12. **Gwynn R, McQuistion HL, McVeigh KH, Garg RK, Frieden TR, Thorpe LE.** Prevalence, diagnosis, and treatment of depression and generalized anxiety disorder in a diverse urban community. Psychiatr. Serv. 2008;59:641–7.
13. **Lorant V, Deliège D, Eaton W, Robert A, Philippot P, Ansseau M.** Socioeconomic inequalities in depression: a meta-analysis. Am. J. Epidemiol. 2003;157:98–112.
14. **Nejtek V, Brown ES, Khan DA, Moore JJ, Van Wagner J, Perantie DC.** Prevalence of mood disorders and relationship to asthma severity in patients at an inner-city asthma clinic. Ann. Allergy Asthma Immunol. 2001;87:129–33.
15. **Pratt L, Brody DJ.** Depression in the United States household population, 2005–2006. *NCHS Data Brief.* 2008:1–8.
16. **Trupin L, Tonner MC, Yazdany J, et al.** The role of neighborhood and individual socioeconomic status in outcomes of systemic lupus erythematosus. J. Rheumatol. 2008;35:1782–8.
17. **Compton W, Conway KP, Stinson FS, Grant BF.** Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991–1992 and 2001–2002. Am. J. Psychiatry. 2006;163:2141.
18. **Kessler R, McGonagle KA, Zhao S, et al.** Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch. Gen. Psychiatry. 1994;51:8–19.
19. **Battaglia TA, Santana CM, Bak S, et al.** Predictors of timely follow-up after abnormal cancer screening among women seeking care at urban community health centers. Cancer. 2010;116:913–921.
20. **Battaglia TA, Roloff K, Posner MA, et al.** Improving follow-up to abnormal breast cancer screening in an urban population. A patient navigation intervention. Cancer. 2007;109:359–367.
21. **Jones GNAS, Jeffries SK, Scarinci IC, Brantley PJ.** Utilization of medical services and quality of life among low-income patients with generalized anxiety disorder attending primary care clinics. Int J Psychiatry Med. 2001;31:183–98.
22. **Fogarty CTSS, Chetty VK, Culpepper L.** Mental health conditions are associated with increased health care utilization among urban family medicine patients. J. Am. Board Fam. Med. 2008;21:398–407.
23. **Weir H, Thun MJ, Hankey BF, et al.** Annual report to the nation on the status of cancer, 1975–2000, featuring the uses of surveillance data for cancer prevention and control. J. Natl. Cancer Inst. 2003;95:1276–1299.
24. **Kroenke KJ, Jackson JL, Chamberlin J.** Depressive and anxiety disorders in patients presenting with physician complaints: clinical predictors and outcome. Am. J. Med. 1997;103:339–347
25. **Aggarwal AFK, Sato A, Adams-Campbell LL, et al.** Are depressive symptoms associated with cancer screening and cancer stage at diagnosis among postmenopausal women? The Women's Health Initiative observational cohort. J Womens Health. 2008;17:1353–1361.
26. **Patten SBWJ, Lavorato DH, Eliasziw M.** The effect of major depression on participation in preventive health care activities. BMC Public Health. 2009;9:87.
27. **Calleo JSM, Greisinger A, Wehmanen O, et al.** Generalized anxiety disorder in older medical patients: diagnostic recognition, mental health management and service utilization. J. Clin. Psychol. Med. Settings. 2009;16:178–185.
28. **Smith-Bindman R, Miglioretti DL, Lurie N, et al.** Does utilization of screening mammography explain racial and ethnic differences in breast cancer? Ann. Intern. Med. 2006;144:541–553.
29. **Yabroff K, Breen N, Vernon SW, Meissner HI, Freedman AN, Ballard-Barbash R.** What factors are associated with diagnostic follow-up after abnormal mammograms? Findings from a U.S. National Survey. Cancer Epidemiol. Biomarkers Prev. 2004;13:723–732.
30. **Ani CBM, Hindman D, Bell D, et al.** Depression symptomatology and diagnosis: discordance between patients and physicians in primary care settings. MC Fam Pract. 2008;9(1):1.
31. **Simpson SM, Krishnan LL, Kunik ME, Ruiz P.** Racial disparities in diagnosis and treatment of depression: a literature review. Psychiatr. Q. 2007;78(1):3–14.
32. **Harman JS, Schulberg HC, Mulsant BH, Reynolds CF.** The effect of patient and visit characteristics on diagnosis of depression in primary care. *J Fam Practice*. 2001; 50(12).
33. **Rawson NS, Malcolm E, D'Arcy C.** Reliability of the recording of schizophrenia and depressive disorder in the Saskatchewan health care datafiles. Soc Psych Psych Epid. 1997;32(4):191–199.
34. **Hostetter AM, Egeland JA, Endicott J.** Amish Study, II: Consensus diagnoses and reliability results. Ame J Psychiat. 1983;140(1):62–66.
35. **Kessler RC, Berglund P, Demler O, et al.** The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). JAMA. 2003;289(23):3095–3105.
36. **Williams DR, Gonzalez HM, Neighbors H, et al.** Prevalence and distribution of major depressive disorder in african americans, caribbean blacks, and non-hispanic whites: Results from the National Survey of American Life. Arch. Gen. Psychiatry. 2007;64 (3):305–315.
37. **Valenstein M, Ritsema T, Green L, et al.** Targeting quality improvement activities for depression. Implications of using administrative data. J Fam Practice. 2000;49(8):721–728.

Factors associated with disclosure of medical errors by housestaff

Andrea C Kronman,[1,2] Michael Paasche-Orlow,[1] Jay D Orlander[1,3]

► An additional appendix is published online only. To view this file please visit the journal online (http://qualitysafety.bmj.com/content/21/4.toc).

[1]Section of General Internal Medicine, Evans Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA
[2]Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, Massachusetts, USA
[3]Medical Service, VA Boston Healthcare System, Boston, Massachusetts, USA

Correspondence to
Dr Andrea C Kronman, Boston University Medical Center, 801 Massachusetts Ave, Suite 470, Boston, MA 02118, USA;
andrea.kronman@bmc.org

Accepted 16 August 2011
Published Online First 22 September 2011

ABSTRACT

Purpose: Attributes of the organisational culture of residency training programmes may impact patient safety. Training environments are complex, composed of clinical teams, residency programmes, and clinical units. We examined the relationship between residents' perceptions of their training environment and disclosure of or apology for their worst error.

Method: Anonymous, self-administered surveys were distributed to Medicine and Surgery residents at Boston Medical Center in 2005. Surveys asked residents to describe their worst medical error, and to answer selected questions from validated surveys measuring elements of working environments that promote learning from error. Subscales measured the microenvironments of the clinical team, residency programme, and clinical unit. Univariate and bivariate statistical analyses examined relationships between trainee characteristics, their perceived learning environment(s), and their responses to the error.

Results: Out of 109 surveys distributed to residents, 99 surveys were returned (91% overall response rate), two incomplete surveys were excluded, leaving 97: 61% internal medicine, 39% surgery, 59% male residents. While 31% reported apologising for the situation associated with the error, only 17% reported disclosing the error to patients and/or family. More male residents disclosed the error than female residents ($p=0.04$). Surgery residents scored higher on the subscales of safety culture pertaining to the residency programme ($p=0.02$) and managerial commitment to safety ($p=0.05$). Our Medical Culture Summary score was positively associated with disclosure ($p=0.04$) and apology ($p=0.05$).

Conclusion: Factors in the learning environments of residents are associated with responses to medical errors. Organisational safety culture can be measured, and used to evaluate environmental attributes of clinical training that are associated with disclosure of, and apology for, medical error.

INTRODUCTION

Everyone makes mistakes. Over the past decade, the medical profession has started to apply a systems approach to patient safety, recognition that coordination of individual, team, and organisational forces are needed to promote patient safety. Analysis of the root causes of an error can prevent future errors by identifying and correcting problems.[1] However, in order to learn from mistakes and develop safer systems, errors must first be identified and reported.

Unfortunately, many errors are never reported. In one study, merely half of the house officers told their attending physicians about the most serious errors they committed.[2] Underreporting of adverse events is estimated to range from 50% to 96% annually.[1 3 4] Rather than dealing with mistakes constructively by reporting and learning from them, studies indicate that physicians typically respond to their mistakes defensively, blaming the system, other members of the healthcare team, or even the patient.[2 5–7] Possible explanations for underreporting medical errors include fear of litigation acting as a deterrent,[8–10] and the professional medical culture that limit an individual's willingness to discuss error.[5 11]

While elements of professional medical culture are hypothesised to lead to widespread underreporting of medical errors, few studies have elucidated and measured aspects of medical culture that are associated with a failure to disclose, particularly in the learning environments of clinical training programmes. In contrast to medical culture, non-medical industries such as aviation and nuclear safety have traditionally valued a professional 'culture of safety', which facilitates reporting of errors, so that individuals operating in groups within an organisation can learn how to prevent future errors.[12] Medical educators have recently attempted to incorporate system-based thinking into their curriculum, in order to incorporate aspects of a safety culture that, along with enquiry and trust, were previously lacking in

residency settings.[2 7 12] Although there is recognition of the need to create a learning culture of safety for residents,[13] measuring educational culture has proved to be a challenge. In addition to the complexities of quantitatively measuring an abstract concept of 'organisational culture',[14] residents train in multiple environments which are dynamic and divergent: their individual clinical teams (which often rotate), the academic residency training programme, and the clinical institution(s) (usually a hospital) each contributing to a trainee's overall sense of culture.

This study endeavours to measure a trainee's perception of their training environment as it relates to safety, and their response to committing an error. By collecting data on both these elements, we explore the relationship between the culture of the training environment and individual behaviour. Specifically, we hypothesised that house officers would be more likely to disclose and apologise for an error if they rate their clinical team as having an environment in which they can report errors without fear of punishment or rejection; rate their training programme as having positive attitudes about reporting and coping with errors in the workplace; and rate the hospital as having a high level of commitment to patient safety. Secondary aims of our study were to evaluate the association between individual characteristics of clinical trainees and, first, disclosure of a medical error, and second, apology for the error.

METHODS

Study design and survey administration

Anonymous, self-administered surveys were distributed to medicine and surgery residents at Boston Medical Center during educational conferences and department meetings in 2005. The distribution and retrieval procedures of the surveys ensured privacy and anonymity of the residents. The researchers (who were also attending physicians) were blinded as to which residents completed the survey, which ensured no conflict of interest for the researchers if asked to evaluate residents. The residents were assured that their privacy and anonymity would be protected, and that the researchers would remain blinded to their participation status. Specifically, the researchers approached a group of housestaff during an educational conference or meeting, explained the purpose of the survey, and then left the room and building. Residents who choose to participate completed the surveys and returned them in sealed envelopes to a box in the room. Those who chose not to participate returned blank surveys in sealed envelopes. At the end of each conference, a research assistant returned to the room to collect the box with the sealed envelopes. All participants received a $10 honorarium, whether or not they completed the survey. The survey, database, and protocol were de-identified. To further protect participants in the event of an accidental breach of anonymity, a certificate of confidentiality was obtained from the National Institutes of Health. The project was carefully reviewed and approved by the institutional review board.

Survey content

Our survey focused on three levels of environment that had the most face validity of microculture constructs within a resident's learning environment. Questions were selected from three validated surveys of organisational culture, adapted for this study to focus on the organisational environment of housestaff. Since the full survey instruments were deemed too high a respondent burden, the authors carefully considered and then selected items from each survey most relevant to the study. The microenvironment of the immediate clinical team was examined with five of seven questions from the Team Psychological Safety Survey, which assesses the belief that well intentioned actions will not lead to punishment or rejection by the team (see online appendix).[15] The macroenvironment of the residency programme was assessed with 10 of the 37 items from the Error Orientation Questionnaire; the items selected assess attitudes to errors and approaches to coping with errors in the workplace (see online appendix).[16] Perception of hospital management's commitment to patient safety on the clinical unit was assessed with four of 19 items from the Patient Safety Survey (see online appendix).[17] Responses were coded with six-point Likert scales, and summed to derive a total score for each survey. For all three scales (Team Psychological Safety, Residency Programme Error Orientation, and Managerial Safety Commitment), higher scores correlated with more positive aspects of culture. In relation to error, participants were asked to recall the circumstances of and share details regarding their most significant medical or surgical error using open-ended text. In a multiple choice format, participants were specifically asked about the following: consequences for the patient; consequences for the resident; if and to whom they disclosed the error; if they had apologised for the error; and perceived causes for the error. Responses to each question were constructed from the results of a previous survey of residents regarding medical error, conducted by Wu *et* al.[2] In addition, residents were asked to characterise their own level of distress from the error using a 10-point Likert scale.

Analysis

Univariate analysis was used to describe demographics, residency type, reporting rates (to colleagues and friends), apology and disclosure rates (to patients),

emotional responses of residents, types of mistakes, and consequences to and responses of both the residents and patients. Errors were classified from the written responses into one of the following categories: procedural, medical management, laboratory test follow-up, delayed diagnosis, or other/not classifiable. We used χ^2 to evaluate differences between categorical variables and Wilcoxon rank sum methods for the three organisational culture scales. In order to compare the subscales with each other, the raw score was converted to a scaled score, by dividing each raw score by the maximum possible score of each subscale, and multiplying by 33.3. We then calculated an overall Medical Culture Summary score, by summing the three scaled subscores, that is, each subscale contributes one-third of the overall Medical Culture Summary score.

RESULTS

Surveys were distributed to 109 residents and 99 surveys were returned, making an overall response rate of 91%. Two residents' surveys were excluded because they reported no mistake, leaving a final population of 97 residents, 59 (61%) from internal medicine residents and 38 (39%) from surgical residents. There were 57 (59%) male residents, of which 33 (58%) were internal medicine residents and 24 (42%) were surgical residents. Two surgical residents did not report their gender, and were excluded from analyses which included gender. The most significant medical or surgical error that was the focus of residents' responses typically occurred in an inpatient setting and during the first year of training (table 1).

Table 1 Characteristics of residents and setting where error occurred

	Total N=97 N (%)
Gender	
Men	57 (59)
Women	38 (39)
Unknown	2 (2)
Programme	
Medicine	59 (61)
Men	33 (56)
Surgery	38 (39)
Men	24 (63)
Training year	
First year	64 (66)
After first year	32 (33)
Unknown	1 (1)
Setting	
Ward	63 (65)
Ambulatory clinic	3 (3)
Operating room	5 (5)
Intensive care unit	22 (23)
Emergency department	4 (4)

Seventy-five per cent of the residents were extremely distressed by their mistake. While 41 (42%) did not provide an adequate description of their error to be classified, 26 (27%) were classified as medication related, 12 (12%) as procedural, 11 (11%) due to delayed diagnosis, and 9 (9%) due to inadequate follow-up to a laboratory test.

Although 20 (21%) of the involved patients had no reported consequences resulting from the errors, common consequences included delayed treatment for 23 (24%), delayed diagnosis for 22 (23%), prolonged hospital stay for 17 (18%), medical complications for 13 (13%), and death for 13 (13%) patients. The errors resulting in patient death were largely errors involving anticoagulants, potassium balance (either not checking blood work or inadequate management of blood potassium level), or insulin. There were no consequences for 60 (62%) of the residents due to the error, but 30 (30%) reported some form of reprimand, 16 (16%) presented the case at a morbidity and mortality conference (which was reported as a consequence), 6 (6%) reported their work and family life was affected, and 1 (1%) was named in a law suit. The most common attributions for the error reported by residents included being too busy (32, 33%) and inexperience (31, 32%). Many residents also attributed their error to having inadequate knowledge, hesitating before acting, or being too tired (table 2). While 30 (31%) reported apologising for the situation associated with the error, only 17 (18%) reported disclosing the error to patients and/or their family. Five residents both disclosed (29% of those who disclosed) and apologised (17% of those who apologised table 3).

Correlates of disclosure and apology

The disclosure rate was higher among surgery residents (24%) than internal medicine residents (14%), but this difference was not statistically significant (p=0.2). Of the residents who disclosed their error, 32 (33%) reported that it was unsupervised. Three (3%) residents reported being told by their attending not to discuss the error with the patient. Female internal medicine residents were significantly less likely to disclose their worst medical error to patients or their families than their male counterparts (p=0.03). In contrast, more female surgery residents, 7 (58%), apologised for their error compared with male surgery residents 7 (29%), though the difference did not reach statistical significance (p=0.1). Of the 13 errors that resulted in a patient's death, only 3 (23%) of the residents disclosed the error to the patient and/or the patient's family, but 6 (46%) residents apologised to the patient's family. More residents who made errors in medication management (8 of 26, 31%) disclosed their error than those who made

Table 2 Characteristics of worst medical errors*

	Total N=97N (%)
Cause of error (self-report)	
Too busy	32 (33)
Inexperience	31 (32)
Inadequate knowledge	19 (20)
Other	15 (15)
Hesitated too long before acting	14 (14)
Too tired	13 (13)
Inadequate communication	10 (10)
Inadequate supervision	9 (9)
Given erroneous information	9 (9)
Felt overwhelmed	8 (8)
Did not obtain the appropriate follow-up test	4 (4)
Inadequate history	3 (3)
Could not recall knowledge	2 (2)
Type of error (classified by investigator)	
Other/could not be classified	41 (42)
Medical management	26 (27)
Procedural	12 (12)
Delayed diagnosis	11 (11)
Did not check lab test	9 (9)
Consequences of error to patient	
Delayed treatment	23 (24)
Delayed diagnosis	22 (23)
None	20 (21)
Prolonged hospital stay	17 (18)
Medical complications	13 (13)
Death	13 (13)
Temporary symptoms	9 (9)
Required additional medications	9 (9)
Other	8 (8)
Additional imaging tests	6 (6)
Additional blood test	5 (5)
Stressed relationship with medical providers	1 (1)
Consequences to resident	
None	60 (62)
Present at M&M conference	16 (16)
Reprimanded by another resident	16 (16)
Reprimanded by attending	12 (12)
Work life affected	6 (6)
Personal life affected	6 (6)
Other	3 (3)
Reprimanded by administrator	2 (2)
Named in lawsuit	1 (1)
Error disclosure/discussion occurred with:	
Fellow residents	66 (68)
Supervising resident	48 (49)
Supervising attending	43 (44)
Significant other	22 (23)
Patient and/or patient's family	17 (18)
Colleagues at a conference	16 (16)
Relative or friend	8 (8)
No one	6 (6)
Other	3 (3)
If disclosure occurred to patient and/or family, who supervised?	
No one	32 (33)
Attending	18 (19)
Senior resident	14 (14)
Risk management	3 (3)
Other	1 (1)

*Residents could select more than one option.
M&M, morbidity and mortality.

errors with procedures (3 of 12, 25%) or delayed diagnosis (2 of 11, 18%) (p=0.05).

Correlates of organisational culture measures

The Safety Culture Summary score was positively associated with disclosure of medical error to the patient and/or patient's family (p=0.04) and apology for the error (p=0.05). There was a trend of association between disclosure and higher scores on the subscales clinical Team Psychological Safety (p=0.07) and Residency Programme Error Orientation scales (p=0.07), but not for Managerial Safety Commitment (p=0.2). Report of apology to the patient and/or patient's family was not associated with the clinical Team Psychological Safety score (p=1.0) but was positively associated with scores on the Residency Programme Error Orientation (p=0.05) and Managerial Safety Commitment (p=0.01). There were no significant gender differences in scores for each of the subscales as well as the summary measure of safety culture. Surgery residents had higher scores on the Residency Programme Error Orientation (p=0.02) and Managerial Safety Commitment scales (p=0.05) compared with medicine residents, but there was no significant difference between programmes in the Safety Culture Summary score (table 4).

CONCLUSIONS

Only 17% of the residents we surveyed reported disclosing their most significant error to their patient and/or patient's family, and only 31% of the residents reported apologising for their most significant error. Our results suggest that factors in the learning environments of the clinical team and residency programme are associated with error disclosure and apology among residents. Individual factors, such as gender and type of error, also appear to be associated with error disclosure

Table 3 Individual factors associated with disclosure and apology, by gender* and programme

	Yes (%)	p Value†
Disclosure to patient and/or patient's family?		
Total (N=97)	17 (18)	0.04
Men (N=57)	14 (25)	
Women (N=38)	3 (8)	
Programme		0.16
Medicine (N=59)	8 (14)	0.05
Men (N=33)	7 (21)	
Women (N=26)	1 (4)	
Surgery (N=38)	9 (24)	0.40
Men (N=24)	7 (29)	
Women (N=12)	2 (17)	
Apologise to patient and/or patient's family?		
Total (N=97)	30 (31)	0.6
Men (N=57)	17 (30)	
Women (N=38)	13 (34)	
Programme		0.2
Medicine (N=59)	16 (27)	0.5
Men (N=33)	10 (30)	
Women (N=26)	6 (23)	
Surgery (N=38)	14 (39)	0.09
Men (N=24)	7 (29)	
Women (N=14)	7 (58)	

*N=95; two respondents who did not enter gender were excluded from the analysis.
†p Values calculated using the χ^2 test.

and apology, and more residents apologised for the error than disclosed it.

Our findings of discordance between apology and disclosure of medical error are consistent with previous studies exhibiting residents may be more willing to apologise for a bad outcome than to reveal that they played a role in causing the bad outcome, resulting in a 'partial disclosure'.[18 19] Collectively, these findings imply that factors that facilitate apologising for an error may differ from influences that facilitate disclosing an error. These findings are reflected in State laws that distinguish different components of conversations with patients about unanticipated outcomes: 'expression of sympathy' (apology), 'explanation' (disclosure), and 'admission of fault', which does not cleanly translate into either category.[20] Additional explanations for the discordance may include the social context in which the error occurred. For example, apologising for a systemic error that occurred would likely be easier than disclosing personal responsibility for an error, which could have greater legal and professional consequences.[19]

The relationships of gender to our outcome measures are complex. More male residents disclosed error (driven mostly by male internal medicine residents) while more women apologised (driven mostly by female surgical residents). With our small sample size, definitive conclusions about the interactions among gender, specialty and disclosure are difficult to ascertain from our data. Previous studies[2 9] have also demonstrated that individual attributes, such as gender and emotional response to the error, influence the reporting rate of the error. However, in contrast to our results, women in a previous study were *more* likely to discuss their errors with their patients and make constructive changes in their practice.[2] Although the female residents in our study were less likely to disclose their error, the female surgical residents were more likely to apologise, consistent with past reports of greater empathy among female physicians.[21] Our results suggest that there are barriers to disclosure in the learning environments of clinical trainees that affect men and women differently. Further research will be needed to elucidate which barriers to disclosure and apology affect genders differentially. For example, possible barriers to disclosure may be attitudinal— women may feel they have more to lose than men by disclosing in order to be professionally successful, or emotional— women may feel more of a sense of helplessness and loss of control once information is disclosed.

We found surgery residents to have higher scores on the residency programme's Error Orientation Scale and the clinical unit's Managerial Safety Commitment Scale than medicine residents, but not on the Safety Culture Summary score. This is consistent with a previous survey of residents, which found that presentations of errors causing adverse events occurred 18% of the time in internal medicine ground rounds compared with 42% in surgery.[22] The differences between these two fields are likely due to divergent regulatory and cultural factors. Historically, morbidity and mortality rounds have served as a forum where surgeons learn from poor outcomes and aspire to identify their errors,[23] but this tradition is weaker in medicine training programmes.[24] The Accreditation Council for Graduate Medical Education (ACGME) requires that surgery morbidity and mortality conferences present and discuss 'all deaths and complications that occur on a weekly basis'. Historically, there has been no similar requirement for internal medicine.[24] Without a specific requirement to do so, adverse events and errors occurring in the medicine service may not be generally discussed.[21]

Several medicine residencies have developed programmes to address the current ACGME competency on Systems Based Practice,[13] by teaching systems-based thinking using root cause analysis of medical errors,[25] which require residents to develop an awareness of working in multidisciplinary teams to enhance patient safety, and participate in identifying system errors and implementing potential solutions.[13] Several studies have demonstrated the benefits of such educational

Table 4 Safety environment factors associated with disclosure and apology, by gender* and programme

	Team Psychological Safety Median range (0–33.3)	p Value†	Residency Programme Error Orientation Median range (0–33.3)	p Value†	Managerial Safety Commitment Median range (0–33.3)	p Value†	Safety Culture Summary Median range (0–100)	p Value†
Total	26.6		25.0		25.0		76.6	
Gender*		0.6		0.3		0.4		0.3
Men	27.8		25.0		26.4		79.1	
Women	26.6		24.4		25.0		76.0	
Programme		0.2		0.02		0.05		0.2
Medicine	27.8		23.9		23.6		75.2	
Surgery	26.6		26.1		27.8		80.5	
Disclosed		0.07		0.07		0.2		0.04
Yes	28.9		26.6		26.4		81.9	
No	26.6		24.4		25.0		76.0	
Apologised		1.0		0.05		0.01		0.05
Yes	27.8		25.5		26.4		79.6	
No	26.6		23.9		25.0		75.5	

*N=95; two respondents who did not enter gender were excluded from the analysis.
†p Values calculated using the Wilcoxon rank-sum test for medians.

interventions,[26–29] although to our knowledge, no study has attempted to measure changes in learning environment as a result of such interventions. Although such programmes are helpful, current ACGME guidelines do not require training programmes to address a thorough behavioural process of managing medical mistakes: accepting responsibility; discussing with colleagues; disclosing and apologising to patient; conducting an error analysis; and making changes in a practice setting designed to reduce future errors.[1 2 5 6 11 30]

The ability of residents to cope with medical error may be dependent on reassurance and learning opportunities provided by medical colleagues and supervisors.[26] Our findings would support this assertion, given the positive association between our derived Medical Culture Summary score and reporting of disclosure and apology. However the low frequency of disclosure and apology suggests that more work needs to be done within our training programmes to mitigate the negative effects of error to individuals, and gain potential benefits from more thorough processing of errors for individuals and the healthcare system.

There are several issues germane to housestaff and errors that are not addressed by our study. Some believe that an effective apology includes offering some form of reparation for the mistake.[31] We did not examine the issue of reparation. While the literature suggests that resident physicians who accept responsibility for their errors and discuss them are more likely to report other improvements in their medical practice,[2] we did not examine this phenomenon. Furthermore, while disclosure in a timely and appropriate manner may influence a patient's decision to pursue legal action,[1–3] we did not explore the relationship between our findings and legal action. In addition, though a doctor's emotional reaction to an error can last for years[32] and negative emotional responses are associated with increased odds of future self-perceived errors,[33] we did not evaluate emotional reactions or predict the future likelihood of error. Lastly, we did not directly examine the extent to which subjects were trained regarding coping with medical error, and hence could not determine if this training influenced their behaviour.

This study has several limitations. First, residents at only one academic medical centre were surveyed, so the results may not be generalisable. In addition, the residents were surveyed during 2004–2005, so it is possible that these results may no longer be accurate. During the past 6 years, the ACGME training requirements have increased their focus on systems-level thinking and training programmes have increasingly focused on reduction of error. However, the authors feel that the key findings of this study are relevant today. The rate of safety culture change is relatively slow, as demonstrated

by a recent hospital survey administered by the Agency for Healthcare Research and Quality: average composite change in safety culture to change 1% over 1–2 years.[20] An increased focus on reduction in error does not translate directly to increases in individual accountability, apology or disclosure of error. Second, the modest sample size, limits opportunity for multivariable analyses as well as statistical power to detect potential associations. Third, our survey directed residents to consider a single error. We did this to focus respondents' attention on the details of an event that they would remember clearly to gain insight into aspects of organisational culture. This specific error may not be representative of most errors. In addition, as most of these errors occurred during the residents' first year of training, the expectation for disclosure and apology may be different than for the other years of training. However, a prior study that included trainees at our institution suggested that the responsibility of delivering bad news often falls to junior members of the team, including first-year trainees and medical students.[34] Fourth, the scales of organisational culture we used have rarely been used in healthcare settings. As such, the clinical significance of our observed score differences are unclear.

Despite the limitations, we successfully adapted survey tools previously used in a business environment to measure aspects of the learning environment of clinical trainees which are associated with disclosure and apology for medical error. This instrument needs to be validated in other institutions before proving its value as a metric in residency programme safety culture. If validated, such an instrument could be a valuable tool to assess changes in learning environments. Measuring culture and providing such feedback to leadership and staff is one of the safe practices recommended by the National Quality Forum to promote patient safety and reduce medical error.[35]

Measuring culture change requires a multimodal approach, of which this instrument could make a valuable contribution.[36] Since the ability to measure medical culture, and changes to it, is immature,[37] our study provides baseline measurements to help move the field further along. Developing learning environment metrics will be valuable to other institutions and training programmes in the coming years, as incremental programmatic changes in systems-level thinking and disclosure of medical error continue to impact the learning environments of residents.

Our results suggest a need for training programmes to provide trainees with structured, meaningful ways to cope with errors to prevent negative emotional responses, as well as create learning environments that facilitate disclosure of errors. Attention may need to be paid to explicate potential gender-related differences. All this is particularly important if, as a profession, we are to instil proper values, attitudes and responses to the inevitable occurrence of error in the next generation of physicians. As residency programmes incorporate systems-level thinking into residency education for patient safety and error prevention, it is important not to neglect the humanistic and interpersonal consequences of error for providers and patients. In order to do so, we need to develop measurement tools for learning environments. Further research is needed to identify successful environmental attributes that promote disclosure and healthy processing of medical errors.

Acknowledgements The authors wish to thank Victoria Brower, MPH, for creating and maintaining the study database, and Ignacio De La Cruz for help with manuscript preparation.

Funding This work was supported by an American Cancer Society Physician Training Award in Preventive Medicine (PT APM 97 185 04, Jane Liebshutz, Principal Investigator) and a Boston University Medical Center Risk Management Grant (Andrea Kronman, Principal Investigator).

Competing interests None.

Ethics approval Boston University Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. *To Err is Human: Building a Safer Health System*. Washington, DC: National Academy Press, 1999.
2. Wu AW, Folkman S, McPhee SJ, *et al*. Do house officers learn from their mistakes? *JAMA* 1991;265:2089–94.
3. Barach P, Small S. Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. *Br Med J* 2000;320:759–63.
4. Leape LL. Error in medicine *JAMA* 1994;272:1851–7.
5. Goldberg RM. Coping with medical mistakes and errors in judgement. *Ann Emerg Med* 2002;39:287–92.
6. Christensen JF, Levinson W, Dunn PM. The heart of darkness: the impact of perceived mistakes on physicians. *J Gen Intern Med* 1992;7:424–31.
7. Mizrahi T. Managing medical mistakes: idealogy, insularity, and accountability among internists-in-training. *Soc Sci Med* 1984;19:135–46.
8. White AA, Gallagher TH, Krauss MJ, *et al*. The attitudes and experiences of trainees regarding disclosing medical errors to patients. *Acad Med* 2008;83:250–6.
9. Padmore JS, Jaeger J, Riesenberg LA, *et al*. 'Renters' or 'owners'? Residents' perceptions and behaviors regarding error reduction in teaching hospitals: a literature review. *Acad Med* 2009;84:1765–74.
10. Weissman JS, Annas CL, Epstein AM, *et al*. Error reporting and disclosure systems: views from hospital leaders. *JAMA* 2005;293:1359–66.
11. Hilfiker D. Facing our mistakes. *N Engl J Med* 1984;310:118–22.
12. Hoff T, Jameson L, Hannan E, *et al*. A review of the literature examining linkages between organizational factors, medical errors, and patient safety. *Med Care Res Rev* 2004;61:3–37.
13. Hoff TJ, Pohl H, Bartfield J. Creating a learning environment to produce competent residents: the roles of culture and context. *Acad Med* 2004;79:532–9.
14. Don H, Slocum JW Jr. Organizational climate: measures, research and contingencies. *Acad Manag J* 1974;17:255–80.
15. Edmondson A. Psychological safety and learning behavior in work teams. *Adm Sci Q* 1999;44:350–83.
16. Rybowiak V, Garst H, Frese M, *et al*. Error Orientation Questionnaire (EOQ): reliability, validity, and different language equivalence. *J Organ Behav* 1999;20:527–47.
17. Pronovost P, Weast B, Holzmueller C, *et al*. Evaluation of the culture of safety: survey of clinicians and managers in an academic medical center. *Qual Saf Health Care* 2003;12:405–10.

18. Fein SP, Hilborne HL, Spiritus EM, *et al*. The many faces of error disclosure: a common set of elements and a definition. *J Gen Intern Med* 2007;22:755–61.
19. Coffey M, Thomson K, Tallett S, *et al*. Pediatric residents' decision-making around disclosing and reporting adverse events: the importance of social context. *Acad Med* 2010;85:1619–25.
20. Sorra J, Famolaro T, Dyer N, *et al*. Hospital survey on patient safety culture: 2011 user comparative database report. AHRQ Publication No. 11-0030, April 2011. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/qual/hospsurvey11/. Accessed 19 Sep 2011.
21. Kaldjian LC, Jones EW, Rosenthal GE, *et al*. An empirically derived taxonomy of factors affecting physicians' willingness to disclose medical errors. *J Gen Intern Med* 2006;21:942–8.
22. Pierluissi E, Fischer MA, Campbell AR, *et al*. Discussion of medical errors in morbidity and mortality conferences. *JAMA* 2003;290:2838–42.
23. Gore DC. National survey of surgical morbidity and mortality conferences. *Am J Surg* 2006;191:708–14.
24. Orlander JD, Fincke BG. Morbidity and mortality conference: a survey of academic internal medicine departments. *J Gen Intern Med* 2003;18:656–8.
25. Voss JD, May NB, Schorling JB, *et al*. Changing conversations: teaching safety and quality in residency training. *Acad Med* 2008;83:1080–7.
26. Engel KG, Rosenthal M, Sutcliffe KM. Residents' responses to medical error: coping, learning, and change. *Acad Med* 2006;81:86–93.
27. Kroll L, Singleton A, Collier J, *et al*. Learning not to take it seriously: junior doctors' accounts of error. *Med Educ* 2008;42:982–90.
28. Stroud L, McIlroy J, Levinson W. Skills of internal medicine residents in disclosing medical errors: a study using standardized patients. *Acad Med* 2009;84:1803–8.
29. Coyle YM, Mercer SQ, Murphy-Cullen CL, *et al*. Effectiveness of a graduate medical education program for improving medical event reporting attitude and behavior. *Qual Saf Health Care* 2005;14:383–8.
30. Leape LL. Reporting of adverse events. *N Engl J Med* 2002;347:1633–8.
31. Lazare A. Apology in medical practice: an emerging clinical skill. *JAMA* 2006;296:1401–4.
32. Waterman AD, Garbutt J, Hazel E, *et al*. The emotional impact of medical errors on practicing physicians in the United States and Canada. *Jt Comm J Qual Patient Saf* 2007;33:467–76.
33. West CP, Huschka M, Novotny PJ, *et al*. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA* 2006;296:1071–8.
34. Orlander JD, Graeme Fincke B, Hermanns D, *et al*. Medical residents' first clearly remembered experiences of giving bad news. *J Gen Intern Med* 2002;17:825–40.
35. National Quality Forum (NQF). Safe practices for better healthcare—2009 update: a consensus report. Washington, DC: NQF;2009. ISBN 978-1-933875-23-1.
36. Measuring Culture Change: A Literature Review. 2006. Colorado Foundation for Medical Care (CFMC). Publication Number: PM-411-114 CO 2006. http://www.cfmc.org/files/nh/MCC%20Lit%20Review.pdf. Accessed 19 Sep 2011.
37. Paine LA, Rosenstein BJ, Sexton JB, *et al*. Assessing and improving safety culture throughout an academic medical centre: a prospective cohort study. *Qual Saf Health Care* 2010;19:547–54.

Factors associated with disclosure of medical errors by housestaff

Andrea C Kronman, Michael Paasche-Orlow and Jay D Orlander

BMJ Qual Saf 2012 21: 271-278 originally published online September 22, 2011
doi: 10.1136/bmjqs-2011-000084

General Internal Medicine
Boston University School of Medicine
2012 Publications L-Z

1. Langmore S, Krisciunas GP, Miloro KV, Evans SR, **Cheng DM**. Does PEG use cause dysphagia in head and neck cancer patients? Dysphagia. 2012 Jun;27(2):251-9. PMID: 21850606.

2. **Lasser KE**, **Kronman AC**, Cabral H, **Samet JH**. Emergency department use by primary care patients at a safety-net hospital. Arch Intern Med. 2012 Feb;172(3):278-80. PMID: 22332165.

3. **Liebschutz JM,** Rothman EF. Intimate-partner violence: What physicians can do. N Engl J Med. 2012 Nov; 367:2071-73 PMID: 23190220.

4. **Linsky A,** Simon SR. Patients' perceptions of their "most" and "least" important medications: A retrospective cohort study. BMC Res Notes. 2012 Nov;5:619. PMID: 23121715.

5. Litwin AH, Smith BD, Drainoni ML, McKee D, **Gifford AL**, Koppelman E, Christiansen CL, Weinbaum CM, Southern WN. Primary care-based interventions are associated with increases in hepatitis C virus testing for patients at risk. Dig Liver Dis. 2012 Jun;44(6):497-503. PMID: 22342471.

6. Lopes RD, Al-Khatib SM, Wallentin L, Yang H, Ansell J, Bahit MC, De Caterina R, Dorian P, Easton JD, Erol C, Ezekowitz JA, Gersh BJ, Granger CB, Hohnloser SH, Horowitz J, **Hylek EM,** McMurray JJ, Mohan P, Vinereanu D, Alexander JH. Efficacy and safety of apixaban compared with warfarin according to patient risk of stroke and of bleeding in atrial fibrillation: A secondary analysis of a randomised controlled trial. Lancet. 2012 Nov;380(9855):1749-58.PMID: 23036896

7. Lopes RD, Becker RC, Newby LK, Peterson ED, **Hylek EM,** Granger CB, Crowther M, Wang T, Carvalho AC, Berwanger O, Giraldez RR, Feitosa GS, Ribeiro JP, Darze E, Kalil RA, Andrande M, Boas FV, Andrade J, Rocha AT, Harrington RA, Lopes AC, Garcia DA. Highlights from the IV International Symposium of Thrombosis and Anticoagulation (ISTA), October 20-21, 2011, Salvador, Bahia, Brazil. J Thromb Thrombolysis. 2012 Jul;34(1):143-63. PMID: 22427055.

8. **Lunze K**, Hamer DH. Thermal protection of the newborn in resource-limited environments. J Perinatol. 2012 May;32(5):317-24. PMID: 22382859. doi: 10.1038/jp.2012.11.

9. Magnani JW, Moser CB, **Murabito JM**, Nelson KP, Fontes JD, Lubitz SA, Sullivan LM, Ellinor PT, Benjamin EJ. Age of natural menopause and atrial fibrillation: The Framingham Heart Study. Am Heart J. 2012 Apr;163(4):729-34. PMID: 22520541.

10. McGinnis KA, Justice AC, Kraemer KL, **Saitz R**, Bryant KJ, Fiellin DA. Comparing alcohol screening measures among HIV-infected and -uninfected men. Alcohol Clin Exp Res. 2012 Oct 10 PMID: 23050632.

11. McKee MM, **Paasche-Orlow MK.** Health literacy and the disenfranchised: The importance of collaboration between limited english proficiency and health literacy researchers. J Health Commun. 2012;17(sup3):7-12. PMID:23030557.

12. Meltzer EC, Averbuch T, **Samet JH, Saitz R, Jabbar K,** Lloyd-Travaglini C, **Liebschutz JM.** Discrepancy in diagnosis and treatment of post-traumatic stress disorder (PTSD): Treatment for the wrong reason. J Behav Health Serv Res. 2012 Apr;39(2):190-201. PMID: 22076315.

13. Meltzer EC, Rybin D, Meshesha LZ, **Saitz R, Samet JH,** Rubens SL, **Liebschutz JM.** Aberrant drug-related behaviors: unsystematic documentation does not identify prescription drug use disorder. Pain Med. 2012 Nov;13(11):1436-43.PMID: 3057631

14. Migneault JP, **Dedier JJ**, Wright JA, Heeren T, Campbell MK, Morisky DE, Rudd P, **Friedman RH**. A culturally adapted telecommunication system to improve physical activity, diet quality, and medication adherence among hypertensive african-americans: A randomized controlled trial. Ann Behav Med. 2012 Feb;43(1):62-73. PMID: 22246660.

15. Mitchell SE, Sadikova E, Jack BW, **Paasche-Orlow MK**. Health literacy and 30-day postdischarge hospital utilization. J Health Commun. 2012;17(sup3):325-338. PMID: 23030580.

16. Muga R, Sanvisens, **Fuster D**, Tor J, Martinez E, Perez-Hoyos S, Munoz A**.** Unhealthy alcohol use, HIV infection and risk of liver fibrosis in drug users with Hepatitis C. PLoS One. 2012;7(10):e46810. PMID: 23056462.

17. **Murabito JM**, White CC, Kavousi M, Sun YV, et al. Association between chromosome 9p21 variants and the ankle-brachial index identified by a meta-analysis of 21 genome-wide association studies. Circ Cardiovasc Genet. 2012 Feb 1;5(1):100-12. PMID: 22199011.

18. **Murabito JM**, Yuan R, Lunetta KL. The search for longevity and healthy aging genes: Insights from epidemiological studies and samples of long-lived individuals. J Gerontol A Biol Sci Med Sci. 2012 May;67(5):470-9. PMID: 22499766.

19. Niederhauser A, VanDeusen Lukas C, Parker V, Ayello EA, Zulkowski K, **Berlowitz D**. Comprehensive programs for preventing pressure ulcers: A review of the literature. Adv Skin Wound Care. 2012 Apr;25(4):167-88. PMID: 22441049

20. **Pace CA,** Lioznov D, **Cheng DM,** Wakeman SE, Raj A, **Walley AY,** Coleman SM, **Bridden C,** Krupitsky E, **Samet JH.** Sexually transmitted infections among HIV-infected heavy drinkers in St Petersburg, Russia. Int J STD AIDS. 2012 Dec;23(12):853-8. PMID: 3258823.

21. Perkins RB, Jorgensen JR, **McCoy ME**, **Bak SM**, **Battaglia TA**, **Freund KM**. Adherence to conservative management recommendations for abnormal pap test results in adolescents. Obstet Gynecol. 2012 Jun;119(6):1157-63. PMID: 22617580.

22. Prushik SG, Farber A, Gona P, Shrader P, Pencina MJ, D'Agostino RB Sr, **Murabito JM**. Parental intermittent claudication as risk factor for claudication in adults. Am J Cardiol. 2012 Mar 1;109(5):736-41. PMID: 22154319.

23. Raj A, Kidd JD, **Cheng DM,** Coleman S, **Bridden C,** Blokhina EA, Krupitsky E, **Samet JH.** Associations between partner violence perpetration and history of STI among HIV-infected substance using men in Russia. AIDS Care. 2012 Sep 13. PMID: 22971088

24. Raj A, Ko N, **Battaglia TA**, Chabner BA, Moy B. Patient navigation for underserved patients diagnosed with breast cancer. Oncologist. 2012 Jul; 17(8):1027-31. PMID: 22752069.

25. Roetzheim RG, **Freund KM**, Corle DK, Murray DM, Snyder FR, **Kronman AC**, Jean-Pierre P, Raich PC, Holden AE, Darnell JS, Warren-Mears V, Patierno S. Analysis of combined data from heterogeneous study designs: an applied example from the patient navigation research program. Clin Trials. 2012 Apr; 9(2):176-87. PMID: 22273587.

26. **Rose AJ.** Improving the management of warfarin may be easier than we think. Circulation. 2012 Nov 6;126(19):2277-9. PMID: 23027800.

27. **Rose AJ**, **Berlowitz DR**, Miller DR, **Hylek EM**, Ozonoff A, Zhao S, Reisman JI, Ash AS. INR targets and site-level anticoagulation control: Results from the Veterans AffaiRs Study to Improve Anticoagulation (VARIA). J Thromb Haemost. 2012 Apr;10(4):590-5. doi: 10.1111/j.1538-7836.2012.04649.x. PMID: 22288563.

28. **Rose AJ**, Petrakis BA, Callahan P, Mambourg S, Patel D, **Hylek EM**, Bokhour BG. Organizational characteristics of high- and low-performing anticoagulation clinics in the Veterans Health Administration. 2012 Aug;47(4):1541-60. PMID: 22299722. doi: 10.1111/j.1475-6773.2011.01377.x.

29. Rosen AK, Itani KM, Cevasco M, Kaafarani HM, **Hanchate A**, Marlena Shin JD, Shwartz M, Loveland S, Chen Q, **Borzecki A**. Validating the patient safety indicators in the Veterans Health Administration: Do they identify true safety events? Med Care. 2012 Jan;50(1):74-85. PMID: 21993057.

30. Rosen AK, Loveland S, Shin M, Shwartz M, **Hanchate A**, Chen Q, Kaafarani HM, **Borzecki A.** Examining the impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration: The case of readmissions. Med Care. 2013 Jan;51(1):37-44.PMID:23032358.

31. Rothendler JA, **Rose AJ,** Reisman JI, **Berlowitz DR,** Kazis LE. Choices in the use of ICD-9 codes to identify stroke risk factors can affect the apparent population-level risk factor prevalence and distribution of CHADS2 scores. Am J Cardiovasc Dis. 2012;2(3):184-91. PMID: 22937488.

32. **Saitz R.** Evidence-based design: Part of evidence-based medicine? Evid Based Med. 2013 Feb;18(1):1. PMID: 22864369. (Editorial)

33. **Saitz R.** Screening and Brief Intervention (SBI): Has it hit the tipping point? Addict Sci Clin Pract. 2012; 7(1): 14. PMID: 23186284. (Editorial)

34. **Saitz R, Samet J.** Science to improve care for people affected by unhealthy alcohol and other drug use. Addict Sci Clin Pract. 2012;7(1):1. PMID: 22966407.

35. Stolk L, Perry JR, Chasman DI, He C, Mangino M, et al, **Murabito JM**. Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways. Nat Genet. 2012 Jan 22;44(3):260-8. doi: 10.1038/ng.1051. PMID: 22267201.

36. Strully KW, Fowler JH, **Murabito JM**, Benjamin EJ, Levy D, Christakis NA. Aspirin use and cardiovascular events in social networks. Soc Sci Med. 2012 Apr;74(7):1125-9. PMID: 22361089.

37. Trikudanathan S, Pedley A, Massaro JM, Hoffmann U, Seely EW, **Murabito JM,** Fox CS. Association of female reproductive factors with body composition: The Framingham Heart Study. J Clin Endocrinol Metab. 2013 Jan;98(1):236-44. PMID: 23093491

38. Tsai AC, Eisa MA, **Crosby SS**, Sirkin S, Heisler M, Leaning J, Iacopino V. Medical evidence of human rights violations against non-Arabic-speaking civilians in Darfur: A cross-sectional study. PLoS Med. 2012 Apr;9(4):e1001198. PMID: 22509136. PMCID: PMC3317898.

39. **Tsui JI**, **Cheng DM**, Libman H, **Bridden C**, **Samet J**. Hepatitis C virus infection is associated with painful symptoms in HIV-infected adults. AIDS Care. 2012 Jul;24(7):820-7. PMID: 22272656.

40. Volandes AE, Brandeis GH, Davis AD, **Paasche-Orlow MK**, Gillick MR, Chang Y, Walker-Corkery ES, Mann E, Mitchell SL. A randomized controlled trial of a goals-of-care video for elderly patients admitted to skilled nursing facilities. J Palliat Med. 2012 Jul;15(7):805-11. PMID: 22559905.

41. **Walley AY**, **Cheng DM**, Pierce CE, Chen C, Filippell T, **Samet JH**, **Alford DP**. Methadone dose, take home status, and hospital admission among methadone maintenance patients. J Addict Med. 2012 Sep;6(3):186-190. PMID: 22694929.

42. **Walley AY,** Tetrault JM, Friedmann PD. Integration of substance use treatment and medical care: A special issue of JSAT. Subst Abuse Treat. 2012 Dec;43(4):377-81. PMID: 23079197 .

43. Walter S, Mackenbach J, Vokó Z, Lhachimi S, Ikram MA, Uitterlinden AG, Newman AB, **Murabito JM**, Garcia ME, Gudnason V, Tanaka T, Tranah GJ, Wallaschofski H, Kocher T, Launer LJ, Franceschini N, Schipper M, Hofman A, Tiemeier H. Genetic, physiological, and lifestyle predictors of mortality in the general population. Am J Public Health. 2012 Apr;102(4):e3-10. PMID: 22397355.

44. Wassel CL, Lamina C, Nambi V, Coassin S, et al, **Murabito JM**. Genetic determinants of the ankle-brachial index: A meta-analysis of a cardiovascular candidate gene 50K SNP panel in the candidate gene association resource (CARe) consortium. Atherosclerosis. 2012 May;222(1):138-47. PMID: 22361517.

45. Williams ED, Bird D, Forbes AW, Russell A, Ash S, **Friedman R,** Scuffham PA, Oldenburg B. Randomised controlled trial of an automated, interactive telephone intervention (TLC Diabetes) to improve type 2 diabetes management: Baseline findings and six-month outcomes. BMC Public Health. 2012; 602-613. PMID: 22857017.

46. Witt DM, Delate T, Garcia DA, Clark NP, **Hylek EM,** Ageno W, Dentali F, Crowther MA. Risk of thromboembolism, recurrent hemorrhage, and death after warfarin therapy interruption for gastrointestinal tract bleeding. Arch Intern Med. 2012 Oct 22;172(19):1484-91. PMID: 22987143.

47. Woz S, Mitchell S, Hesko C, **Paasche-Orlow M**, Greenwald J, Chetty VK, O'Donnell J, Jack B. Gender as risk factor for 30 days post-discharge hospital utilization: A secondary data analysis. BMJ Open. 2012 Apr 18;2(2): PMID: 22514241.

48. You JJ, Singer DE, Howard PA, Lane DA, Eckman MH, Fang MC, **Hylek EM**, Schulman S, Go AS, Hughes M, Spencer FA, Manning WJ, Halperin JL, Lip GY. Antithrombotic therapy for atrial fibrillation: Antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest. 2012 Feb;141(2 Suppl):e531S-75S. PMID: 22315271.

49. Yi PH, Kim JS, Hussein KI, **Saitz R.** Behavior, knowledge, and attitudes towards khat among Yemeni medical students and effects of a seminar. Subst Abus. 2012;33(4):373-7. PMID:22989281.

50. Zallman L, Rubens SL, **Saitz R**, **Samet JH**, Lloyd-Travaglini C, **Liebschutz J**. Attitudinal barriers to analgesic use among patients with substance use disorders. Pain Res Treat. 2012;2012:167062. PMID: 22685649.

Dysphagia
DOI 10.1007/s00455-011-9360-2

ORIGINAL ARTICLE

Does PEG Use Cause Dysphagia in Head and Neck Cancer Patients?

Susan Langmore · Gintas P. Krisciunas · Keri Vasquez Miloro · Steven R. Evans · Debbie M. Cheng

Received: 14 February 2011 / Accepted: 19 July 2011

Abstract Percutaneous endoscopic gastrostomy (PEG) use is common in patients who undergo radiotherapy (RT) for head and neck cancer to maintain weight and nutrition during treatment. However, the true effect of PEG use on weight maintenance and its potential impact on long-term dysphagia outcomes have not been adequately studied. This retrospective study looked at swallowing-related outcomes among patients who received prophylactic PEG vs. those who did not, and among patients who maintained oral diets vs. partial oral diets vs. those who were nil per os (NPO). Outcomes were assessed at the end of RT and at 3, 6, and 12 months post RT. A comprehensive review of patients' medical charts for a 6-year period yielded 59 subjects with complete data. Results showed no difference in long-term percent weight change between the prophylactic PEG patients vs. all others, or between patients who, during RT, had oral diets vs. partial oral diets vs. NPO. However, those who did not receive prophylactic PEGs and those who maintained an oral or a partial oral diet during RT had significantly better diet outcomes at all times post RT. Dependence on a PEG may lead to adverse swallowing ability in post-irradiated head and neck cancer patients possibly due to decreased use of the swallowing musculature.

Keywords Deglutition · Deglutition disorders · Head and neck cancer · Radiotherapy · PEG · Weight · Diet

Dysphagia is a well-known complication of radiotherapy (RT) for head and neck cancer (HNC). The incidence of dysphagia varies greatly by report, with figures ranging anywhere from 9 to 77% [1, 2]. These disparate figures are undoubtedly a result of multiple variables that may impact swallow function such as time after radiation treatment, stage of cancer, additional surgery or chemotherapy, patient characteristics, and response of the patient to the treatment. While instrumental procedures (i.e., fluoroscopy or fiberoptic endoscopic examination of swallowing [FEES]) can yield an objective diagnosis, more functional and more common indicators of a swallowing problem are weight loss, diet level, and the need for a feeding tube.

The need for a feeding tube in patients undergoing radiotherapy is a particularly interesting variable since it is viewed as both beneficial and detrimental for patients. In cancer clinics across the U.S., different institutional policies have prevailed. Some centers avoid placing a feeding tube unless the patient is showing signs of extreme weight loss or reported inability to swallow. Other centers place feeding tubes prophylactically in all patients (or those with advanced-stage cancer) before the RT regime even begins. Evidence for one policy over the other is lacking. Instead,

This work was done entirely at Boston University Medical Center.

S. Langmore (✉) · G. P. Krisciunas
Otolaryngology Department, Boston University Medical Center, 820 Harrison Ave., FGH Bldg, 4th Floor, Boston, MA 02118, USA
e-mail: langmore@bu.edu

S. Langmore · K. V. Miloro
Boston University Sargent College of Health and Rehabilitation Sciences, 635 Commonwealth Ave., Boston, MA 02215, USA

S. R. Evans
Data Coordinating Center, Boston University School of Public Health, 801 Massachusetts Ave., Boston, MA 02118, USA

D. M. Cheng
Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Ave., Boston, MA 02118, USA

Published online: 18 August 2011

institutional precedent and anecdotal experience are more influential in guiding a physician in his recommendation.

Evidence for best practice in this area is scant. The American Head and Neck Cancer Society's guidelines regarding PEGs state that "Due to the location of their tumors and the mucosal effects of therapy, many patients with head and neck cancer are nutritionally compromised. If highly toxic therapy is planned, placement of a PEG tube prior to starting therapy enables patients to maintain nutrition during therapy and recover more quickly" [3]. These recommendations, which appear to favor prophylactic PEG placement, come from a consensus of practitioners who were selected by the Society and are not necessarily based on evidence from well-designed research studies.

In contrast to these guidelines, recent awareness of the significant dysphagia experienced by some patients after RT and the emphasis on swallowing rehabilitation have led some clinicians to believe that relying on a feeding tube during RT may adversely affect a patient's ultimate ability to swallow. This belief originates from the perception that if a person has a feeding tube during his course of RT, he will depend on it for nutrition and will stop taking anything by mouth because of pain, altered taste, and lack of appetite. While refraining from oral intake may be more comfortable for the patient, clinicians fear that the associated reduction in swallowing frequency may ultimately lead to muscle atrophy and augment the severity of radiation fibrosis in the throat. This belief is not based on strong research either, but it is suggested from one pilot study that showed some benefit to patients who exercised their swallowing muscles during RT [4].

The best evidence for guiding clinical practice comes from well-designed prospective or retrospective research studies. Ideally, patients who received a prophylactic PEG (placed before or during the first week of RT) would be compared to patients who received a therapeutic PEG (placed after the first week of RT) or to patients who never received a PEG and maintained a 100% oral diet. The outcome of interest would be swallowing status after a given period of time.

A literature search for all such studies published between January 1991 and January 2010 yielded only ten studies with acceptable evidence, including two prospective randomized control trials. Four studies compared prophylactic vs. therapeutic PEG placement [5–8], four studies compared prophylactic PEG to oral/no PEG patients [9, 10], and two studies compared prophylactic PEG to therapeutic PEG and to oral patients (or pooled therapeutic and oral patients) [11, 12]. None of these studies reported swallow status from an instrumental swallow examination; instead, they reported weight loss, use of a feeding tube, or patient/clinician report as surrogate indicators of swallow function.

Looking at *weight loss* as a possible measure of swallow function, the literature suggested that prophylactic PEG favors less weight loss over the course of RT treatment [5, 9, 10, 12–14]. Only one study, albeit a randomized clinical trial, found no significant difference in BMI change at the end of RT between patients who received prophylactic PEG vs. patients with therapeutic PEG or oral diets [11]. Reports of comparative weight loss for long-term outcomes were mixed. One study found no significant difference between the prophylactic PEG, therapeutic PEG, and 100% oral diet groups at 3 or 6 months post RT [14]. Two other studies found that prophylactic PEG patients fared better (lost less weight) than therapeutic PEG patients at 12 months post RT [7, 12].

Diet levels were not reviewed in any of the studies, but two studies did look at swallow function via clinician/patient rating scales. Chen et al. [14] found that significantly more patients who had been given a prophylactic PEG reported dysphagia at 3 and 6 months compared to patients who had 100% oral diets. Morton et al. [7] found that patients who had a PEG placed for longer durations, regardless of prophylactic or therapeutic nature, reported worse swallowing function at 12 months ($P < 0.01$).

In summary, there is a paucity of published evidence regarding dysphagia outcomes, and adequate guidelines for PEG placement are clearly lacking. In light of the preliminary results of Chen and Morton suggesting PEGs may detrimentally affect long-term dysphagia outcomes, the present study was undertaken to determine if PEG use influenced long-term swallow function in HNC patients who received radiotherapy at our institution. The first aim was to determine whether patients who received a prophylactic PEG had better or worse swallowing function, diet scores, or weight change outcomes than patients who did not receive a prophylactic PEG. The second aim of this study was to determine whether patients who used a feeding tube for all, some, or none of their nutrition fared differently in terms of eventual swallowing ability, diet level, and weight change. The results of the second question would hopefully cast light on the question of whether the act of eating and using the muscles involved in swallowing would impact long-term dysphagia outcomes.

Materials and Methods

A retrospective chart review was conducted of all head and neck cancer patients treated with RT from January 2003 to September 2008 at Boston University Medical Center. The chart reviews were performed by three authors (KVM, SEL, GPK) using the medical center electronic database and the clinical data warehouse. An initial search of various hospital databases revealed 880 patient records. After

excluding patients with ineligible cancer locations (e.g., esophagus, skin) or who did not fall within our eligibility window, a final list of 150 patients remained.

Patients were then excluded from analysis if there was incomplete outcome data at baseline, incomplete outcome data at more than one time point of interest, if they received only a partial dose of RT, or if they had persistence or recurrence of HNC at the time of outcome. Of the 150 medical records, there were 59 patients who met all inclusion and exclusion criteria, and for whom adequate data was available up to 1 year post RT. Approval from the Institutional Review Board of Boston University Medical Center was obtained prior to data collection.

Age, gender, location and stage of cancer, chemotherapy, and extent of surgery were recorded for each patient. Weight, diet level, and PEG status were noted at five time periods: start of RT, completion of RT, 3 months post RT, 6 months post RT, and 12 months post RT. Diet level at each time point was categorized by 6 diet levels listed in Table 1.

Objective instrumental evaluation of swallowing ability would have been an excellent indicator of function, but very few patients had formal swallow studies so this could not be considered. Accordingly, our primary outcomes of interest were percent weight loss and mean diet levels over time, both representing common surrogates of swallow function and dysphagia.

Percent weight loss, rather than absolute weight loss, was used to help account for vastly different body sizes. While weight at the *start of RT* was probably not the subject's usual weight, it was a good baseline from which change over time could be compared, especially to capture the effect of PEG use on weight during RT. However, *end of RT* was chosen as the baseline when recording *diet level* so that changes to this outcome would be a result of variables other than the acute effects of RT (such as pain). The end of RT represented a nadir in diet that was experienced by almost all patients, and any tumor-related dysphagia symptoms had been resolved. So this end-of-RT nadir served as a "normalized" starting point from which we could record changes in diet level over time between different groups of patients.

Table 1 Diet level categories

Diet level	Description
1	NPO
2	Partial PEG/partial oral
3	Liquid diet
4	Puree and wet soft foods
5	Edentulous or regular food with difficulty or requiring liquid wash
6	Regular diet with no alterations needed

Data Analysis

Two sets of analyses were conducted to determine the impact of PEG use on long-term diet level and weight change. All analyses used the same cohort of patients, although they were categorized differently depending on the question of interest.

The first set of analyses evaluated whether patients who received a prophylactic PEG (main independent variable) had different outcomes (%weight change and diet scores) than people who did not receive a prophylactic PEG. The second set of analyses assessed whether diet level at the end of RT for all patients (main independent variable) was associated with the two outcomes of interest. Outcomes of interest for both analyses were percent weight change and diet score at 3, 6, and 12 months after completion of RT.

For each of the above analyses, generalized linear mixed-effects models were used to evaluate the relationships of interest. The mixed-effects models were used to account for the correlation from using repeated observations on the same subject. Covariates considered were chemotherapy, stage of cancer, and major surgery. However, due to insufficient data in some cells, the results were adjusted only for surgery. Potential collinearity of covariates was confirmed by calculating the correlation between independent variables; no pair of variables had a Spearman correlation > 0.40. The association between diet status at baseline and diet scores at follow-up was evaluated in preliminary analyses using the nonparametric Kruskal–Wallis test at each time point, and the results were consistent with the parametric regression analyses. For the diet level analysis, the Tukey–Kramer method was used to identify significant pairwise differences across the different baseline diet groups. Analyses were performed using SAS software ver. 9.1 (SAS Institute, Cary, NC).

Results

Analysis 1: Prophylactic PEG vs. No/Therapeutic PEG

Table 2 summarizes patient descriptive characteristics in this analysis. Of the 59 patients in our cohort, 33 had a PEG placed at some time over the first year. The vast majority (27/33) of patients who had a PEG received it prophylactically (before or during the first week of RT). For this first set of analyses, these 27 patients were our "prophylactic PEG" cohort. Six of the 33 patients had a PEG placed therapeutically: 3 patients had a PEG placed during weeks

Table 2 Descriptive characteristics for prophylactic PEG vs. no/therapeutic PEG analysis

Characteristics	All patients ($n = 59$)	Prophylactic PEG ($n = 27$)	No PEG/therapeutic PEG ($n = 32$)
Age (years)			
Mean	58.8	57.1	60.3
Range	35–80	40–72	35–80
Sex			
Male	41 (69)[d]	18 (67)	23 (72)
Female	18 (31)	9 (33)	9 (28)
PEG placement	–	27 prophylactic	26 never 6 therapeutic
Disease stage[a]			
I	3 (5)	–	3 (9)
II	7 (12)	–	7 (22)
III	11 (19)	7 (26)	4 (13)
IV	37 (64)	20 (74)	12 (53)
Not staged	1 (2)	–	1 (3)
Tumor site			
Oral cavity, oropharynx, and nasopharynx	38 (64)	21 (78)	17 (53)
Hypopharynx and larynx	13 (22)	3 (11)	10 (31)
Malignant neoplasm of other and ill-defined sites	8 (14)	3 (11)	5 (16)
Surgery			
Minor[b]	31 (52)	15 (55)	16 (50)
Major[c]	11 (19)	4 (15)	7 (22)
None	17 (29)	8 (30)	9 (28)
Chemotherapy	47 (80)	25 (93)	22 (69)
Type of RT			
IMRT	26 (44)	12 (44)	14 (44)
3D/conventional/other	32 (54)	14 (52)	18 (56)
Unknown	1 (2)	1 (4)	–

Except for age, all values are number (percent)

[a] Stage according to AJCC TNM Staging Guide, 2004

[b] Biopsy, neck dissection, and/or tonsillectomy

[c] Any surgery more extensive than biopsy, neck dissection, and/or tonsillectomy

[d] Values in parentheses indicate percent of group that the adjacent number represents

2–6 of RT, 2 patients had it placed 6–8 months after RT completion, and 1 patient's date of PEG placement was not known. Our second cohort included these six therapeutic PEG patients as well as the 26 patients who never received a feeding tube.

Median duration of PEG use was 97 days or about 8 months (range = 61–2,049 days). Over the year, many patients had their PEG removed while a few had a PEG placed. The overall number of patients with a feeding tube was (1) 27/59 or 46% of all patients at the end of RT; (2) 21/59 or 36% at 3 months after the end of RT (20 were prophylactic); (3) 14/59 or 24% at 6 months (12 prophylactic); and (4) 9/59 or 15% at 12 months (7 were prophylactic).

There were 41 men and 18 women in the sample, with a mean age of 58.8 years. The most common sites for cancer were the oral cavity, oropharynx, and nasopharynx (78% in the prophylactic PEG group; 50% in the no/therapeutic PEG group). The majority of patients in both groups had advanced stage IV cancer and had chemotherapy added to their treatment, with a higher proportion in the prophylactic PEG group for each condition. Rates of minor, major, or no surgery were similar in both groups.

In order to determine whether the patients who received a prophylactic PEG had different outcomes over time than patients who did not receive a prophylactic PEG, percent weight loss and mean diet scores were analyzed over time. Table 3 summarizes these results.

Percent Weight Change

Analyses were performed using linear mixed-effects models adjusted for surgery. Percent weight loss between the two groups was analyzed from baseline at the start of RT to the end of RT, and 3, 6, and 12 months post RT. The differences were not statistically significant across time ($P = 0.19$).

Table 3 Adjusted mean differences in percent weight change and diet levels based on prophylactic PEG use (yes: $n = 27$, no: $n = 31$) at completion of RT

	% weight change		Diet level	
	Adjusted mean difference (95% CI)[a]	*P* value	Adjusted mean difference (95% CI)[a]	*P* value
Prophylactic PEG (yes vs. no)	−2.19 (−5.50, 1.12)	0.19	−1.11 (−1.78, −0.44)	0.002

[a] Adjusted for surgery and time since baseline

Change in Diet Level

Analyses explored whether prophylactic PEG placement affected diet status across time (from baseline at the end of RT to 3, 6, and 12 months after RT). In contrast to weight change results, the prophylactic PEG group had significantly worse diet outcomes than either the no PEG group or the therapeutic PEG group across time ($P = 0.002$).

In summary, these two analyses showed that prophylactic PEG placement by itself did not significantly influence long-term weight changes in head and neck cancer patients, but it did appear to affect diet level up to 1 year after RT. In order to probe these findings further, the next set of analyses were conducted.

Analysis II: Outcomes by Type of Nutritional Intake: 100% PEG Diet vs. Partial PEG/Partial Oral Diet vs. 100% Oral Diet

In this patient population, diet status and feeding tube status are not always analogous. This is because some patients who have a PEG still eat a full oral diet while others are partially or completely reliant on their feeding tubes. Since some people with PEGs had consistent oral diets, we speculated that PEG placement by itself may not be the most important factor influencing long-term diet outcomes. Instead, we wondered if PEG dependence, and hence less frequent use of swallowing musculature, may be more important in influencing long-term dysphagia.

Accordingly, in this second analysis we divided our 59 subjects into three cohorts based on each individual's diet level at the end of RT, regardless of PEG status. These three groups were (1) 100% NPO (nil per os) (diet level 1) representing very little swallowing muscle use, (2) partial PEG/partial oral (diet level 2) representing intermediate swallowing muscle use, and (3) 100% oral intake (diet levels 3–6) representing most swallowing muscle use. As in the first analysis, we analyzed percent weight loss and mean diet scores between groups at baseline and 3, 6, and 12 months after RT. Our hypothesis was that consistent eating and swallowing during RT would help preserve a patient's ability to swallow and thus lead to a more normal long-term diet.

Table 4 delineates descriptive characteristics in this second analysis. At the end of RT, 14 patients were 100% PEG dependent (forming group 1), 8 patients had a partial PEG and partial oral diet (forming group 2), and 37 patients had a 100% oral diet (forming group 3). Of the 37 patients who had a 100% oral diet, 9 had PEGs but were not using them. Overall, most patients reported only minor changes in diet prior to RT (average score was 4.8/6). By the end of RT, however, many subject's diet levels dropped dramatically, with the average diet score falling to 3.4/6 (liquid or puree/wet soft food).

Patients who relied partially or completely on their PEGs for nutrition all had advanced stage III or IV cancer, while 71% of patients who had a 100% oral diet had stage 3 or 4 cancer. Across all three groups, oral cavity, oropharynx, and nasopharynx cancers were most common, and minor surgery was a common adjuvant therapy. The distribution of major, minor, and no surgery was similar in all groups. The addition of chemo agents to the radiation therapy was common, ranging from 73 to 100%. Intensity-modulated radiation therapy (IMRT) and 3D/conventional radiation therapy techniques were equally represented among the groups.

Percent Weight Change

Percent weight change was recorded at the start of RT, the end of RT, and at 3, 6, and 12 months post RT. Figure 1 shows the mean change in weight over time in these three diet groups ($n = 59$). Longitudinal regression analyses adjusted for surgery revealed no significant difference in weight change across the three groups during the 1 year follow-up (see Table 5). Thus, the results were similar to the earlier analysis of weight change between prophylactic PEG users and all other patients.

Change in Diet Level

Since the data suggested that PEG placement or PEG use alone may not significantly affect weight change, a major question of interest was whether the subjects who were completely dependent on their feeding tubes (FTs) at the end of RT were able to advance their long-term diet levels

Table 4 Descriptive characteristics of total PEG vs. partial PEG vs. 100% oral groups at baseline (end of RT)

Characteristics	All patients ($n = 59$)	Group 1: all intake via PEG ($n = 14$)	Group 2: partial PEG/ partial oral ($n = 8$)	Group 3: 100% oral ($n = 37$)
Age (years)				
Mean	58.8	61.3	54.9	58.8
Range	35–80	51–79	51–68	35–80
Sex				
Male	41 (69)[d]	9 (64)	7 (88)	25 (68)
Female	18 (31)	5 (36)	1 (12)	12 (32)
Disease stage[a]				
I	3 (5)	–	–	3 (8)
II	7 (12)	–	–	7 (19)
III	11 (19)	5 (36)	1 (12)	5 (14)
IV	37 (63)	9 (64)	7 (88)	21 (57)
Not staged	1(2)	–	–	1 (3)
Tumor site				
Oral cavity, oropharynx, and nasopharynx	38 (64)	10 (71)	5 (62)	23 (62)
Hypopharynx and larynx	13 (22)	2 (14)	3 (38)	8 (22)
Malignant neoplasm of other and ill-defined sites	8 (14)	2 (14)	–	6 (16)
Surgery				
Minor[b]	31 (52)	11 (79)	4 (50)	16 (43)
Major[c]	11 (19)	–	3 (38)	8 (22)
None	17 (29)	3 (21)	1 (12)	13 (35)
Chemotherapy	47 (80)	12 (86)	8 (100)	27 (73)
Type of RT				
IMRT	26 (44)	5 (36)	4 (50)	17 (46)
3D/conventional/other	32 (54)	8 (57)	4 (50)	20 (54)
Unknown	1 (2)	1 (7)	–	–

Except for age, all values are number (percent)

[a] Stage according to AJCC TNM Staging Guide, 2004

[b] Biopsy, neck dissection, and/or tonsillectomy

[c] Any surgery more extensive than biopsy, neck dissection, and/or tonsillectomy

[d] Values in parentheses indicate percent of group that the adjacent number represented

as well as subjects who were partially dependent on their FTs or as well as those who had 100% oral diets. Using the same three groups defined by diet level at the end of RT, we analyzed mean group diet scores at 3, 6, and 12 months post RT. Preliminary Kruskal–Wallis tests were used to analyze these data (separately by time point), and indicated significant differences in mean diet score between the three groups at all time points ($P < 0.01$). Linear mixed-effects models were then used to evaluate differences in average diet score across all groups longitudinally after adjusting for surgery. Once again, there was a significant difference across all three groups (global $P < 0.001$).

The post-hoc Tukey–Kramer multiple-comparisons procedure showed that both the partial PEG/partial oral group and the 100% oral group had significantly higher mean diet scores than the 100% PEG group over time. There were no significant differences in mean diet scores between the partial PEG/partial oral group and the 100% oral group. Results of statistical analyses are presented in Table 5 and are represented graphically in Fig. 2.

Discussion

The motivation for the current study was the possibility that extended PEG use might hinder swallowing ability. Two indicators of adverse swallowing outcomes were investigated: percent weight loss and diet level. The results of this study showed that there was no significant difference in weight loss among all groups at any time after radiation therapy. It also found that patients who had PEGs placed prophylactically, or patients who relied totally on their PEGs for nutrition, had significantly worse diet outcomes than patients who did not receive prophylactic PEGs or who continued to eat orally despite receiving a PEG.

Two recent studies suggested that patients who have prophylactic PEGs or who use a PEG for an extended period of time experience adverse outcomes [7, 14]. However, this study is the first to present objective evidence that these patients may develop a more severe dysphagia than patients who do not receive prophylactic PEGs or who do not rely on a PEG for all their nutrition.

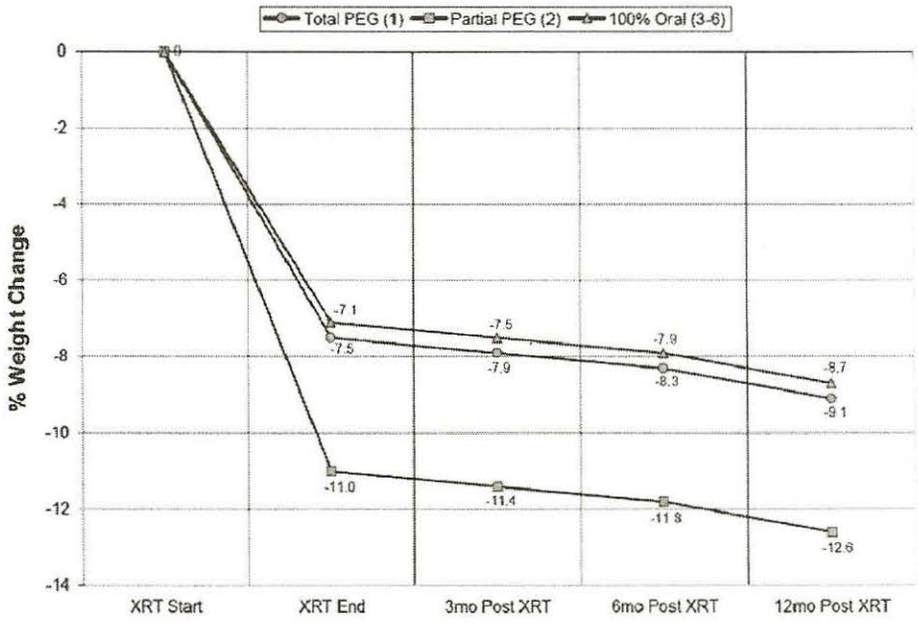


Fig. 1 Mean percent weight change of three groups from start of RT to 1 year post RT (unadjusted at RT start, adjusted for surgery at all other time points). Groups were categorized by diet level at completion of RT (baseline)

Table 5 Adjusted mean differences in percent weight change and diet levels based on diet level[a] at completion of RT

	% weight change		Diet level	
	Adjusted mean difference (95% CI)[b]	*P* value	Adjusted mean difference (95% CI)[b]	*P* value
Diet level[c]		0.27		<0.001
Group 2 vs. 1	−3.56 (−10.50, 3.37)		1.61 (0.53, 2.70)[d]	
Group 3 vs. 1	0.39 (−4.55, 5.33)		2.49 (1.73, 3.26)[d]	
Group 3 vs. 2	3.95 (−1.85, 9.76)		0.88 (−0.03, 1.79)	

[a] Group 1: 100% PEG, $n = 14$; Group 2: partial PEG/partial oral, $n = 8$; Group 3: 100% oral, $n = 37$

[b] Adjusted for surgery and time since baseline

[c] Reported *P* values for predictor correspond to global test

[d] Significantly different based on Tukey–Kramer method

Placing PEG tubes prophylactically before radiotherapy is a common practice in many institutions across the U.S. This study suggests that there may be real risks involved in this practice and it should be reanalyzed for its risk/benefit ratio. The median duration of PEG placement in the group studied here was 8 months, which agrees closely with the average duration of prophylactic PEG placement reported in the literature [15–17]. This is a substantial amount of time that a patient must endure an artificial means of feeding, not to mention the associated discomfort, stigma, and hindrance to social eating. The need for long-term PEG use follows a very difficult and traumatic cancer treatment and its associated morbidities and serves to exacerbate an already-decreased quality of life.

The most common reason for placing a PEG prophylactically is to maintain weight, hydration, and nutrition, especially during the painful weeks of radiotherapy which can make this treatment more endurable and in some instances safer. Surprisingly, the current study did not find a significant difference in percent weight change between prophylactic PEG and no/therapeutic PEG patients during RT, in contrast to several other published studies mentioned earlier [5, 9, 10, 12–14]. The current study also found no significant differences in weight loss between the prophylactic PEG and no/therapeutic PEG patients at any time point after RT. When the groups were reorganized into three groups according to diet level at the end of RT rather than the presence or the absence of a prophylactic feeding tube, we still found no significant differences in percent weight loss among the groups.

While long-term weight outcomes at 3, 6, or 12 months after RT are not often reported in the literature, the current findings were in agreement with the one other study that did investigate this outcome [14]. The reason for the discrepancy between findings from the current study and most of the literature is unclear. One may question whether

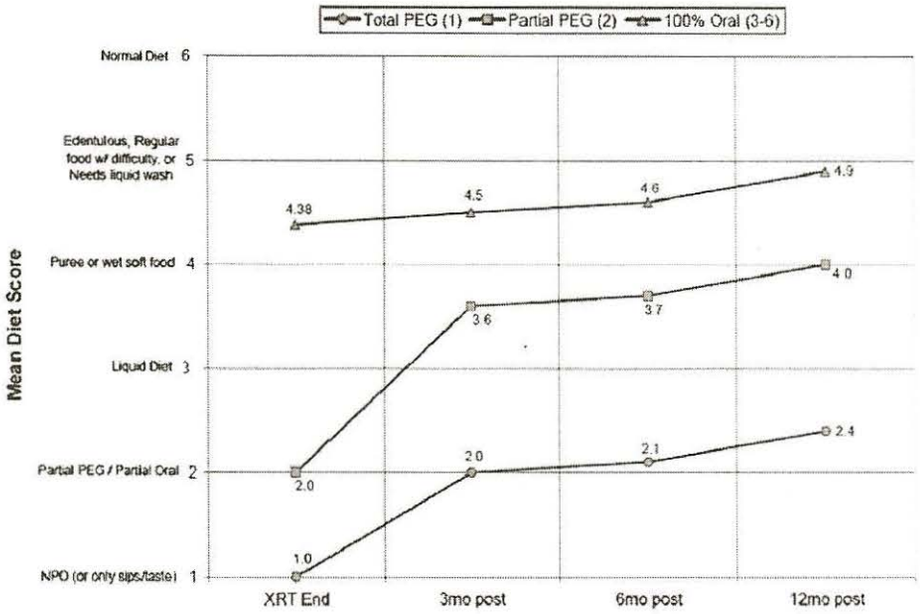


Fig. 2 Mean diet score over time in three groups defined by diet level at baseline (unadjusted at RT end, adjusted for surgery at all other time points)

confounding factors such as chemotherapy, stage of cancer, or any other patient or treatment variables could have accounted for our percent weight change results. It may be that the prophylactic PEG and the 100% PEG use groups had a greater share of these adverse factors that would have caused a significantly greater decrease in weight had they not received a prophylactic PEG and/or relied on their PEG. Such a conclusion would suggest that prophylactic PEGs and PEG reliance actually helped these patients minimize weight loss during RT as well as the other patients who did not have these other risk factors.

Table 2, which summarizes the patient demographics, does not support this hypothesis. The majority of patients in all groups had stage III/IV cancer (68% of those patients without prophylactic PEG) and received chemotherapy (74% of that same group). However, while the groups looked fairly similar demographically, they were not equally matched and the limited sample size precluded us from adjusting statistically for any confounding variable except surgery. Therefore, this question remains unsettled. A future study addressing this paradoxical issue would be extremely interesting and incredibly useful in guiding PEG placement policies.

Diet level was the other outcome investigated in this study. This is arguably the most functional and meaningful indicator of swallowing ability. Change in diet over the first year after RT has not been reported systematically in the literature, but it turned out to be very revealing in this study. When the patients were identified and grouped by the diet they were taking rather than whether they had a PEG, it became possible to study the effect on swallowing outcomes of eating and drinking orally versus depending on a PEG. The motivating question was whether consistent use of the swallowing musculature during and after RT determined long-term diet status. Would PEG reliance, perhaps coupled with radiation-induced fibrosis, translate into disuse atrophy significant enough to promote a chronic dysphagia?

This analysis found that (1) the 100% oral group had significantly higher mean diet scores than the 100% PEG group across time, (2) the partial PEG/partial oral group had significantly higher mean diet scores than the 100% PEG group across time, but (3) there were no significant differences in mean diet scores between the partial PEG/partial oral and 100% oral groups. These were surprising findings. First, we showed that patients who were 100% dependent on their PEGs for nutrition (and remained NPO) not only presented with worse diet outcomes over the first year compared to patients who consumed 100% of their diets orally, but also had worse diet outcomes than patients who had a PEG but continued to take some nutrition orally. Most of those 100% PEG users had their PEGs removed sometime over the course of the year, but their diet levels did not catch up to those of the other patients. In Fig. 2, it can be seen that the 100% PEG intake group (as defined by their diet status at end of RT) were on average consuming only an oral liquid diet at 6 and 12 months post RT, whereas the partial PEG and 100% oral groups were on average consuming several food consistencies. This very functional outcome is clearly meaningful to a person's daily life. Second, this analysis showed that patients who continued to take some nutrition orally, in spite of having a PEG, attained the same long-term diet outcomes as patients who never had a PEG. This suggests that continued oral intake during RT allows for continued use of the swallowing musculature, which in turn may prevent the

degree of dysphagia experienced by other patients who rely entirely on their PEGs.

The implications of this study are twofold. First, the community and the individual physician should reconsider recommending prophylactic PEG placement in patients who are likely to depend entirely on the PEG for nutrition unless it is absolutely necessary. There will always be patients who benefit from a PEG and they certainly should not be denied its convenience and ability to deliver much needed nutrition. Those patients probably consist of individuals who are severely malnourished, fragile, and have shown that they are not able to prevent a sharp and consistent weight loss.

Second, and perhaps more important, when a patient is given a PEG, this study suggests that they should be strongly encouraged to continue swallowing something orally on a regular basis. If the purpose is to prevent atrophy and loss of function of the swallowing musculature, it does not matter what the bolus is: the patient can simply sip water throughout the day. The point is to continue exercising and moving the muscles so that they do not become atrophied, stiff, and potentially locked in by fibrotic tissue.

This interpretation is still hypothetical and should be confirmed by larger studies that are able to control for all known confounders. A prospective randomized clinical trial would provide the best evidence, of course, and would mitigate the limitations of the current study. Since prophylactic PEG placement is still the standard of care in many institutions, an RCT could feasibly be carried out whereby all prophylactic PEG patients are assigned to one of two groups, where one remains NPO and the other takes some nutrition orally throughout the duration of their PEG use. A major improvement over the current study, and all others that have addressed this topic, would be to include formal swallow studies, a standardized diet score, and a QOL scale. The issue is too important to let the status quo stand unchallenged.

Acknowledgment This work was supported by a grant from the NIH/NCI #RO1CA120950-03.

References

1. Cooper JS, Fu K, Marks J, Silverman S. Late effects of radiation therapy in the head and neck region. Int J Radiat Oncol Biol Phys. 1995;31(5):1141–64.
2. List MA, Bilir SP. Evaluations of quality of life and organ function. Semin Oncol. 2004;31(6):827–35.
3. National comprehensive cancer network clinical practice guidelines in oncology, 2008. Available at http://www.nccn.org/professionals/physician_gls/f_guidelines.asp (accessed 2010).
4. Carroll WR, Locher JL, Canon CL, Bohannon IA, McColloch NL, Magnuson JS. Pretreatment swallowing exercises improve swallow function after chemoradiation. Laryngoscope. 2008; 118(1):39–43.
5. Beer KT, Krause KB, Zuercher T, Stanga Z. Early percutaneous endoscopic gastrostomy insertion maintains nutritional state in patients with aerodigestive tract cancer. Nutr Cancer. 2005; 52(1):29–34.
6. Lee NY, de Arruda FF, Puri DR, Wolden SL, Narayana A, Mechalakos J, et al. A comparison of intensity-modulated radiation therapy and concomitant boost radiotherapy in the setting of concurrent chemotherapy for locally advanced oropharyngeal carcinoma. Int J Radiat Oncol Biol Phys. 2006;66(4):966–74.
7. Morton RP, Crowder VL, Mawdsley R, Ong E, Izzard M. Elective gastrostomy, nutritional status and quality of life in advanced head and neck cancer patients receiving chemoradiotherapy. ANZ J Surg. 2009;79(10):713–8.
8. McLaughlin BT, Gokhale AS, Shuai Y, Diacopoulos J, Carrau R, Heron DE, et al. Management of patients treated with chemoradiotherapy for head and neck cancer without prophylactic feeding tubes: the University of Pittsburgh experience. Laryngoscope. 2010;120(1):71–5.
9. Hearne BE, Dunaj JM, Daly JM, Strong EW, Vikram B, LePorte BJ et al. Enteral nutrition support in head and neck cancer: tube vs. oral feeding during radiation therapy. J Am Diet Assoc. 1985;85(6):669, 674, 677.
10. Mercuri A, Lim Joon D, Wada M, Rolfo A, Khoo V. The effect of an intensive nutritional program on daily set-up variations and radiotherapy planning margins of head and neck cancer patients. J Med Imaging Radiat Oncol. 2009;53(5):500–5.
11. Salas S, Baumstarck-Barrau K, Alfonsi M, Digue L, Bagarry D, Feham N, et al. Impact of the prophylactic gastrostomy for unresectable squamous cell head and neck carcinomas treated with radio-chemotherapy on quality of life: prospective randomized trial. Radiother Oncol. 2009;93(3):503–9.
12. Beaver ME, Matheny KE, Roberts DB, Myers JN. Predictors of weight loss during radiation therapy. Otolaryngol Head Neck Surg. 2001;125(6):645–8.
13. Chang JH, Gosling T, Larsen J, Powell S, Scanlon R, Chander S. Prophylactic gastrostomy tubes for patients receiving radical radiotherapy for head and neck cancers: a retrospective review. J Med Imaging Radiat Oncol. 2009;53(5):494–9.
14. Chen AM, Li B, Lau DH, Farwell DG, Lu Q, Stuart K, et al. Evaluating the role of prophylactic gastrostomy tube placement prior to definitive chemoradiotherapy for head and neck cancer. Int J Radiat Oncol Biol Phys. 2010;78(4):1026–32.
15. Nguyen NP, North D, Smith HJ, Dutta S, Alfieri A, Karlsson U, et al. Safety and effectiveness of prophylactic gastrostomy tubes for head and neck cancer patients undergoing chemoradiation. Surg Oncol. 2006;15(4):199–203.
16. Foster JM, Filocamo P, Nava H, Schiff M, Hicks W, Rigual N, et al. The introducer technique is the optimal method for placing percutaneous endoscopic gastrostomy tubes in head and neck cancer patients. Surg Endosc. 2007;21(6):897–901.
17. Mekhail TM, Adelstein DJ, Rybicki LA, Larto MA, Saxton JP, Lavertu P. Enteral nutrition during the treatment of head and neck carcinoma: is a percutaneous endoscopic gastrostomy tube preferable to a nasogastric tube? Cancer. 2001;91(9):1785–90.

Susan Langmore PhD

Gintas P. Krisciunas MS, MPH

Keri Vasquez Miloro MS

Steven R. Evans MPH

Debbie M. Cheng ScD

expenditure for Part A and B services for normal-weight beneficiaries was $6832, but expenditures were significantly less for overweight and obese beneficiaries ($5473 [$P<.001$] and $5790 [$P=.02$], respectively). Expenditures increased over time for all groups but appeared to increase more quickly in obese beneficiaries.

In regression analysis adjusting for demographic and economic covariates, expenditures increased by a mean of $122 per year ($P<.001$) for normal-weight beneficiaries, and increased significantly faster for overweight (excess increase of $108 per year [$P=.01$]) and obese beneficiaries (excess increase of $149 per year [$P=.001$]). Adjusting for chronic conditions accounted for differences in the trend across BMI groups. After adjusting for chronic conditions, interactions between overweight and time ($P=.71$) and obesity and time ($P=.98$) were no longer significant.

Comment. Although Medicare expenditures increased in all BMI groups over this period, expenditures increased significantly faster for overweight and obese Medicare beneficiaries. Increasing rates of weight-related chronic conditions over time appeared to account for this trend.

We found smaller obesity-related differences in expenditures than reported in previous research.[1,5,6] This may be due to differences in expenditure data (claims vs estimated expenditures), costs included (Medicare only vs total costs), or participant age range.

Our results suggest that projections related to the future costs of obesity should take into account changes in chronic health conditions among the obese older population as drivers of increased expenditures. If the parallel trends of increasing obesity and increasing numbers of chronic conditions continue, obesity-related Medicare spending may rise faster than projected based on prevalence of obesity alone.

Dawn Alley, PhD
Jennifer Lloyd, MA
Thomas Shaffer, MHS
Bruce Stuart, PhD

Author Affiliations: Department of Epidemiology and Public Health, University of Maryland School of Medicine (Dr Alley and Ms Lloyd), Doctoral Program in Gerontology, University of Maryland (Ms Lloyd and Mr Shaffer), and Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy (Dr Stuart, Ms Lloyd, and Mr Shaffer), Baltimore.
Correspondence: Dr Alley, Department of Epidemiology and Public Health, University of Maryland, 660 W Redwood St, Ste 221B, Baltimore, MD 21201 (dalley@epi.umaryland.edu).
Author Contributions: Mr Shaffer had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Alley and Stuart. *Acquisition of data:* Shaffer and Stuart. *Analysis and interpretation of data:* Alley, Lloyd, Shaffer, and Stuart. *Drafting of the manuscript:* Alley, Lloyd, and Stuart. *Critical revision of the manuscript for important intellectual content:* Shaffer and Stuart. *Statistical analysis:* Shaffer. *Obtained funding:* Stuart. *Administrative, technical, and material support:* Lloyd. *Study supervision:* Alley.
Financial Disclosure: None reported.
Funding/Support: This study was supported by the Commonwealth Foundation and grants K12 HD043489 and T32 AG000262 from the National Institutes of Health.
Role of the Sponsors: Funding agencies were not involved in study design, collection, management, analysis, or interpretation of data, or manuscript preparation.

1. Lakdawalla DN, Goldman DP, Shang B. The health and cost consequences of obesity among the future elderly. *Health Aff (Millwood)*. 2005;24(suppl 2): W5R30-W5R41.
2. Yang Z, Hall AG. The financial burden of overweight and obesity among elderly Americans: the dynamics of weight, longevity, and health care cost. *Health Serv Res*. 2008;43(3):849-868.
3. Flegal KM, Graubard BI, Williamson DF, Gail MH. Cause-specific excess deaths associated with underweight, overweight, and obesity. *JAMA*. 2007;298(17): 2028-2037.
4. Alley DE, Chang VW. The changing relationship of obesity and disability, 1988-2004. *JAMA*. 2007;298(17):2020-2027.
5. Thorpe KE, Florence CS, Howard DH, Joski P. The impact of obesity on rising medical spending. *Health Aff (Millwood)*. 2004(suppl Web exclusives):W4-480-W4-486.
6. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Aff (Millwood)*. 2009;28(5):w822-w831.
7. Sing M, Banthin JS, Selden TM, Cowan CA, Keehan SP. Reconciling medical expenditure estimates from the MEPS and NHEA, 2002. *Health Care Financ Rev*. 2006;28(1):25-40.
8. Cawley J, Burkhauser RV. *Beyond BMI: the Value of More Accurate Measures of Fatness and Obesity in Social Science Research*. Cambridge, MA: National Bureau of Economic Research; 2006.

Emergency Department Use by Primary Care Patients at a Safety-Net Hospital

In fee-for-service payment models, there are strong financial incentives for hospitals to tolerate high levels of emergency department (ED) use, including use by established primary care patients. Yet, as health care reform introduces global payment models, high levels of ED use will no longer be financially tenable. Understanding the magnitude of the problem of ED use by established primary care patients is crucial to redesigning primary care delivery and reimbursement in the United States. We had 2 objectives: (1) to characterize ED use at an urban safety-net hospital after the implementation of Massachusetts health reform, focusing on patients who had primary care providers (PCPs) and (2) to identify patterns of ED use that might inform the hospital-based primary care practices' transformation to a medical home, and eventually, to an accountable care organization.

Methods. Boston Medical Center (BMC) is an urban safety-net hospital with 8 primary care practices staffed by 105 PCPs. The practices predominantly serve a minority and low-income population. We identified patients who had 1 or more primary care visits from July 1, 2009, to July 1, 2010, and examined their ED use over this period. We defined frequent ED users as patients with 4 or more ED visits in the past year and occasional ED users as those with 1 to 3 ED visits in the past year. We

Table. Demographic, Clinical, and Utilization Characteristics of Primary Care Patients According to ED Use, 2009-2010

Variable %	No ED Visits (n = 25 888)	Occasional ED Use[a] (n = 11 787)	Frequent ED Use[b] (n = 1928)
Age, mean (SD), y	45.8 (17.3)[f]	46.9 (17.2)	46.8 (16.8)
Female sex	57.9	58.2[g]	51.7
Language			
English	78.6[f]	78.0[g]	87.0
Spanish	5.8	7.7	7.1
Haitian Creole	7.2	6.6	2.3
Other	8.4	7.7	3.5
Race[c]			
White	34.2[f]	16.0[g]	16.2
Black/African American	44.2	61.6	63.5
Hispanic/Latino	11.1	15.0	17.0
Other	10.5	7.4	3.3
Insurance			
Medicare	15.0[f]	18.3[g]	28.6
Commercial	32.6	17.7	9.4
Medicaid	16.8	24.5	33.6
Free care	8.6	13.2	8.3
Commonwealth care[d]	10.4	11.8	12.0
Other	16.6	14.4	8.1
Medical comorbidity[e]			
COPD	4.0[f]	7.1[g]	14.3
Diabetes	16.8[f]	23.6[g]	30.0
CHF	1.7[f]	4.9[g]	10.8
Psychiatric comorbidity[e]			
Anxiety	10.9[f]	13.8[g]	23.4
Bipolar disorder	1.8[f]	3.4[g]	9.0
Depression	20.2[f]	29.5[g]	47.1
Posttraumatic stress disorder	4.0[f]	8.3[g]	18.1
Panic disorder	1.3[f]	1.9[g]	3.8
Schizophrenia	1.0[f]	1.7[g]	4.0
Substance use diagnoses[e]			
Any alcohol	3.8[f]	7.2[g]	17.0
Alcohol dependence	1.0[f]	2.5[g]	10.7
Any drug	3.6[f]	6.3[g]	14.8
ED visits in the past year, mean (SD), No.	0.0	1.5 (0.7)[g]	6.4 (5.5)
Severity of ED visits			
Low	NA	75.4[g]	72.2
Intermediate	NA	17.5[h]	19.0
High	NA	7.1[g]	8.8
Primary care visits in the past year, mean (SD), No.	2.7 (2.2)[f]	3.5 (2.8)[g]	4.3 (3.5)

Abbreviations: CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; ED, emergency department; NA, not applicable.
[a] Emergency department visits in the past year, 1 to 3.
[b] Emergency department visits in the past year, 4 or more.
[c] Patient race and ethnicity were determined by clinical registration staff.
[d] Commonwealth Care is a Massachusetts insurance program for poor and near-poor uninsured adults.
[e] Diagnoses were obtained from outpatient medical record problem list.
[f] $P < .001$ (compared with ≥1 ED visit).
[g] $P < .001$ (compared with frequent ED use).
[h] $P < .01$ (compared with frequent ED use).

used an algorithm developed by Billings et al[1] to categorize each visit's principal *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) diagnosis and to determine the probability that a visit required ED care. Using a validation of this algorithm,[2] we defined visits as high severity if the probability that ED care was needed was 0.75 or higher for the visit's principal *ICD-9-CM* diagnosis. Similarly, we defined visits as low severity if the probability that ED care was needed was 0.25 or lower. Visits of indeterminate severity were defined as those with a probability higher than 0.25 and lower than 0.75. Visits classified as high severity have been found to have a strong association with future hospitalization or death.[3] We performed χ^2 tests and *t* tests to compare differences in demographics between persons with and without any ED use and with occasional vs frequent ED use. The Boston University Medical Campus institutional review board has approved this study as exempt.

Results. During the study period (2009-2010), 39 603 patients had 1 or more primary care visits. Most patients (65.4%) did not make any ED visits, while 11 787 (29.8%) were occasional ED users and 1928 were frequent ED users (4.9%) (**Table**). The 11 787 patients with occasional

ED use made 17 759 visits over the study period, while the 1928 patients with frequent ED use made 12 289 visits. Approximately half (49.8%) of all ED visits occurred on weekdays, while BMC primary care practices were open. Most ED visits were for low-severity conditions.

Comment. Emergency department use by primary care patients at an urban safety-net hospital was high, though most visits were of low severity. One possible reason for this is lack of access to primary care,[4] with few available appointments to see a PCP. While data on time to third next available appointment, a standard measure of primary care access,[5] are not available for the primary care practices during the study period, other practice metrics suggest that access may have been a problem. For example, missed primary care appointment rates were high, averaging 24.5%. High missed appointment rates are often correlated with long wait times to schedule appointments.[6] In addition, monthly telephone call statistics show that only between 72.4% and 88.1% of patient telephone calls were answered by the primary care call center over the study period. It is possible that patients called the practices with an urgent problem, did not have their telephone call answered promptly, and decided to seek care in the ED instead. Indeed, 13% of telephone calls were abandoned by patients over the study period (patients called and subsequently hung up while they were kept on hold). The fact that nearly half of all ED visits took place during the hours of primary care clinic operation further suggests that appointment availability may have been an issue. In addition, a sizable minority, roughly one-fifth, of primary care is provided by residents,[7] who have limited availability when they are not in clinic. It is also possible that Massachusetts health reform has affected access to primary care. As newly insured patients have entered primary care in large numbers, it is possible that access to primary care has worsened for other patients.

Massachusetts has been a bellwether for the implementation of health reform and will be a bellwether for the transformation of primary care, with the move away from fee-for-service payments and the introduction of global payments for health care. Overall ED volume has continued to increase in Massachusetts after health reform.[8] It is unclear if changes in primary care practice and payment will be sufficient to reduce high levels of ED use among patients at an urban safety-net hospital.

Karen E. Lasser, MD, MPH
Andrea C. Kronman, MD, MSc
Howard Cabral, PhD, MPH
Jeffrey H. Samet, MD, MA, MPH

Author Affiliations: Section of General Internal Medicine, Department of Medicine, Boston Medical Center and Boston University School of Medicine (Drs Lasser, Kronman, and Samet), and Departments of Community Health Sciences (Drs Lasser and Samet), Biostatistics (Drs Cabral), and Epidemiology (Dr Samet), Boston University School of Public Health, Boston, Massachusetts.
Correspondence: Dr Lasser, Health/Care Disparities Research Program, 801 Massachusetts Ave, Boston Medical Center, Section of General Internal Medicine, Room #2091, Boston, MA 02118 (karen.lasser@bmc.org).
Author Contributions: Dr Lasser, the primary author, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Lasser, Kronman, and Samet. *Acquisition of data:* Lasser. *Analysis and interpretation of data:* Lasser, Kronman, Cabral, and Samet. *Drafting of the manuscript:* Lasser, Cabral, and Samet. *Critical revision of the manuscript for important intellectual content:* Lasser, Kronman, Cabral, and Samet. *Statistical analysis:* Lasser, Kronman, and Cabral. *Obtained funding:* Lasser and Samet. *Administrative, technical, and material support:* Lasser. *Study supervision:* Lasser and Samet.
Financial Disclosure: Dr Lasser was a consultant to Rise Health (June-November 2011).
Funding/Support: This study was supported by the Department of Medicine at Boston Medical Center/Boston University School of Medicine.
Previous Presentation: This work was presented at the Society of General Internal Medicine Annual Meeting; May 7, 2011; Phoenix, Arizona.
Additional Contributions: Meredith Manze D'Amore, PhD (Section of General Internal Medicine, Boston Medical Center/Boston University School of Medicine), assisted in manuscript preparation; Linda C. Rosen, BSEE (Boston University Office of Clinical Research), performed data extraction; and Maxim D Shrayer, PhD (Department of Slavic and Eastern Languages and Literatures, Boston College), and William Fernandez, MD (Department of Emergency Medicine, Boston Medical Center/Boston University School of Medicine), provided comments on an earlier draft of the manuscript.

1. Billings J, Parikh N, Mijanovich T. Emergency department use: the New York Story. *Issue Brief (Commonw Fund)*. 2000;(434):1-12.
2. Ballard DW, Price M, Fung V, et al. Validation of an algorithm for categorizing the severity of hospital emergency department visits. *Med Care*. 2010; 48(1):58-63.
3. Wharam JF, Landon BE, Galbraith AA, Kleinman KP, Soumerai SB, Ross-Degnan D. Emergency department use and subsequent hospitalizations among members of a high-deductible health plan. *JAMA*. 2007;297(10):1093-1102.
4. Smulowitz P, Landon BE, Burke L, Baugh C, Gunn H, Lipton R. Emergency department use by the uninsured after health care reform in Massachusetts. *Intern Emerg Med*. 2009;4(6):501-506.
5. Murray M, Berwick DM. Advanced access: reducing waiting and delays in primary care. *JAMA*. 2003;289(8):1035-1040.
6. Lasser KE, Mintzer IL, Lambert A, Cabral H, Bor DH. Missed appointment rates in primary care: the importance of site of care. *J Health Care Poor Underserved*. 2005;16(3):475-486.
7. Zallman L, Ma J, Xiao L, Lasser KE. Quality of US primary care delivered by resident and staff physicians. *J Gen Intern Med*. 2010;25(11):1193-1197.
8. Smulowitz PB, Lipton R, Wharam JF, et al. Emergency department utilization after the implementation of Massachusetts health reform. *Ann Emerg Med*. 2011;58(3):225-234.e1.

Proton Beam Therapy and Treatment for Localized Prostate Cancer: If You Build It, They Will Come

The number of treatment options for localized prostate cancer continues to expand, amidst growing concern regarding overdiagnosis and overtreatment of low-risk disease.[1-3] Treatment patterns, however, may be driven by availability of novel tech-

Intimate-Partner Violence — What Physicians Can Do

Jane M. Liebschutz, M.D., M.P.H., and Emily F. Rothman, Sc.D.

The U.S. Centers for Disease Control and Prevention (CDC) recently released a comprehensive report on the prevalence of sexual violence, stalking, and intimate-partner violence (IPV) in the United States.[1] The report relays the alarming findings that 35.6% of women in this country are raped, assaulted, or stalked by intimate partners at some point during their lives, and approximately 6% experience these events in any given year. Men are also at risk for IPV victimization: 28.5% report having been victimized at some time during their lifetime, and 5% report victimization within the past year. But the forms and consequences of IPV experienced by women and men are not the same. Women are more than twice as likely as men to experience sexual coercion in their intimate relationships (17% vs. 8%) and are twice as likely to experience severe forms of physical assault by an intimate partner, such as being choked, hit with a fist, or kicked (24.3% vs. 13.8%). The most striking differences relate to the consequences: very few men (5.2%) report ever being fearful of their intimate partners, in contrast to 28.8% of women, and women are almost four times as likely as men to be injured by a partner (14.8% vs. 4.0%).

The costs of IPV are burdensome, for the health care system and for society. A decade ago, the CDC estimated the cost of IPV to the United States to be $5.8 billion per year ($10.4 billion in 2012 dollars), and it's been estimated that the cost of providing health care to adult survivors of IPV ranges from $2.3 billion to $7.0 billion in the first year after the assault. The annual health care costs for women who are experiencing ongoing IPV are 42% higher than those for nonabused women. This finding is unsurprising, given the evidence that IPV victimization of women increases the risks of injury, gastrointestinal disorders, chronic pain, central nervous system symptoms (including fainting and seizures), hypertension, and gynecologic problems.[2]

What can physicians do about IPV? All health care providers should be alert to aspects of patients' histories or symptoms that could suggest IPV and then should follow up with specific questions. According to the U.S. Preventive Services Task Force, screening asymptomatic female patients for IPV victimization may provide benefits, with minimal adverse effects.[3] As of August 2012, new guidelines under the Affordable Care Act require insurance coverage to include IPV screening and counseling as part of eight essential health services for women at no additional cost to the patient.[4] Therefore, at a minimum, all primary care physicians should now be screening female patients 12 years of age or older for IPV. Specialty professional organizations recommend that obstetricians and pediatricians also consider performing regular IPV screening. Numerous IPV screening instruments may be used to begin a dialogue with the patient; one of them (known as HITS) is shown in the table.[5] Another question that may be used to start a discussion about safety at home is simply, "Are you afraid of your partner or anyone else?"

There are several steps doctors should take when patients report potential IPV. First, clinicians should acknowledge the patient's admission of abuse: we advise thanking the patient for trusting the provider with the information and expressing concern about the patient's safety. Second, we suggest asking the patient if he or she would like to be connected to IPV advocacy services. If patients do want legal assistance, counseling, shelter, or other services, local domestic violence agencies affiliated with the state coalition are likely to be the most reliable resources (see box). Third, clinicians should offer the patient the National Domestic Vio-

The HITS Screening Tool for Domestic Violence.*

How Often Does Your Partner	Never	Rarely	Sometimes	Fairly Often	Frequently
Physically hurt you	1	2	3	4	5
Insult or talk down to you	1	2	3	4	5
Threaten you with harm	1	2	3	4	5
Scream or curse at you	1	2	3	4	5

* A total score of more than 10 is suggestive of intimate partner violence. This information, called R3, is available as a free Android or iPhone app. From Sherin et al.[5]

Key Resources

National Domestic Violence Hotline, www.thehotline.org/, 1-800-799-SAFE (7233). Provides crisis intervention, information, and referrals for victims of domestic violence.

Futures without Violence, www.futureswithoutviolence.org, a national organization dedicated to improving the health care response to violence; offers resources and information for providers and health care organizations.

Virtual Lecture Hall, www.vlh.com/domesticviolenceCME. Evidence-based online IPV training, free access for 30 days, $25/credit hour for CME certificate.

National Network to End Domestic Violence, www.nnedv.org. National organization of state domestic violence coalitions.

Child Welfare Information Gateway, Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services, www.childwelfare.gov. Information and resources on child abuse and neglect.

lence hotline number (see box); the hotline makes printed, pocket-size handouts (palm cards) available to providers who wish to distribute them to patients. Fourth, clinicians should consider whether child protective services are required. In many states, the abuse of one parent by another does not necessitate a report to child protective services, so it's up to the clinician to determine whether a report is warranted. Clinicians should consider inviting the patient to make the report directly in order to increase the likelihood that staff members at child protective services agencies will view the patient as able to maintain a safe household for the children. Fifth, they should screen the patient for coexisting depression, anxiety, and substance abuse. And they should use caution when prescribing sedatives, since the sedating action may diminish patients' physical or mental ability to defend themselves or deescalate tensions.

When patients screen negative for IPV but the provider nevertheless suspects that they're experiencing abuse, it's important that the provider not force disclosure. It's not critical that the patient acknowledge IPV victimization in order to benefit from the screening. Asking IPV-related questions signals to the patient that the provider is caring and concerned, trustworthy, and willing to discuss the topic during a future visit. Moreover, simply being asked the questions may prompt the patient to reconsider privately whether his or her relationship is healthy. And of course providers need not receive a positive screening response in order to provide universal education about IPV. Even if a patient screens negative, we would encourage the provider to state that many patients do experience IPV at some point and that there are many resources to help people who feel unsafe in their relationships. Handing palm cards with the national hotline number to all patients and encouraging them to take one for a friend if they wish, for example, may be an effective way of providing help to victims who don't feel comfortable disclosing their situations.

There are several ways in which providers can do unintentional harm to patients who are experiencing IPV. Asking no questions about IPV may signal that the provider is not a potential resource for the patient. But the manner in which IPV victimization is documented in patient records can have ramifications for child custody cases. Detailed information about best practices for IPV documentation is available from the national organization Futures without Violence (see box). In addition, providers should refrain from telling patients who are experiencing IPV what they must do (e.g., "you need to leave"). Only trained experts in IPV advocacy are qualified to help victims determine their own best course to safety. There is a potential for lethal and injurious harm, particularly when one partner attempts to leave the relationship. For this reason, actively ensuring that the link between the patient and an IPV advocacy agency is made successfully is the best practice. Finally, it's critically important that providers respect the confidentiality of patients who are experiencing IPV. Not only do victims face stigma and prejudice, but employers and insurers could potentially discriminate against them if their status became known.

IPV is now recognized as a substantial public health problem. Health care providers can play a critical role in helping to reduce and prevent IPV by screening and referring patients to appropriate resources, familiarizing themselves with best practices related to IPV documentation and victim response, and presenting themselves as caring and trustworthy allies for their patients who are experiencing abuse. Research has established that health care–based screenings and interventions for IPV can benefit patients,[3] and the Affordable Care Act ensures that preventive care will include these screenings for women and ado-

lescents. There is thus some cause for hope that we may curb the violence and play a role in creating safer homes and safer families nationwide.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

From Boston University School of Medicine (J.M.L.) and the Boston University School of Public Health (E.F.R.) — both in Boston.

1. Black M, Basile K, Breiding M, et al. The National Intimate Partner and Sexual Violence Survey: 2010 summary report. Atlanta: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 2011.

2. Campbell JC. Health consequences of intimate partner violence. Lancet 2002;359:1331-6.

3. Nelson HD, Bougatsos C, Blazina I. Screening women for intimate partner violence: a systematic review to update the U.S. Preventive Services Task Force Recommendation. Ann Intern Med 2012;156:796-808.

4. James L, Shaeffer S. Interpersonal and domestic violence screening and counseling: understanding new federal rules and providing resources for health providers. Futures without violence. May 2012 (http://www.futureswithoutviolence.org/userfiles/file/HealthCare/FWV-screening_memo_Final.pdf).

5. Sherin KM, Sinacore JM, Li XQ, Zitter RE, Shakil A. HITS: a short domestic violence screening tool for use in a family practice setting. Fam Med 1998;30:508-12.

DOI: 10.1056/NEJMp1204278

BMC Research Notes

SHORT REPORT Open Access

Patients' perceptions of their "most" and "least" important medications: a retrospective cohort study

Amy Linsky[1,2,3,4,5*] and Steven R Simon[1,2,3,4,5]

Abstract

Background: Despite benefits of adherence, little is known about the degree to which patients will express their perceptions of medications as more or less important to take as prescribed. We determined the frequency with which Veteran patients would explicitly identify one of their medications as "most important" or "least important."

Findings: We conducted a retrospective cohort study of patients from ambulatory clinics at VA Boston from April 2010-July 2011. Patients answered two questions: "Which one of your medicines, if any, do you think is the most important? (if none, please write 'none')" and "Which one of your medicines, if any, do you think is the least important? (if none, please write 'none')." We determined the prevalence of response categories for each question. Our cohort of 104 patients was predominantly male (95%), with a mean of 9 medications (SD 5.7). Regarding their most important medication, 41 patients (39%) identified one specific medication; 26 (25%) selected more than one; 21 (20%) wrote "none"; and 16 (15%) did not answer the question. For their least important medication, 31 Veterans (30%) chose one specific medication; two (2%) chose more than one; 51 (49%) wrote "none"; and 20 (19%) did not directly answer the question.

Conclusions: Thirty-five percent of patients did not identify a most important medication, and 68% did not identify a least important medication. Better understanding of how patients prioritize medications and how best to elicit this information will improve patient-provider communication, which may in turn lead to better adherence.

Keywords: Communication, Adherence, Veterans, Quality of care, Patient safety

Findings

Background and objectives

Medication adherence may improve clinical outcomes, but approximately half of all prescriptions are not taken as prescribed [1-3]. Patient medication-taking behavior is influenced by many factors, including health literacy, socioeconomic status, perceived medication necessity, future health concerns and whether the drug provides symptom relief [4-6]. Further, patients' beliefs about their medications are dynamic and can fluctuate with changes in symptoms, competing health- and non-health-related demands and trust in the health care provider [5,7].

Patient non-adherence appears to imply that some degree of prioritization of medications is occurring, although not necessarily in an explicit manner. Moreover, it is unclear to what degree patients will express their perceptions of medications as more or less important to their treating health care provider and whether clinicians concur with patients' prioritization schemas. To begin to address these questions, we sought to determine the frequency with which Veteran patients would explicitly identify one of their medications as "most important" or "least important," as well as to characterize the medications selected.

Methods

We conducted a retrospective cohort study of a convenience sample of patients from ambulatory care clinics

* Correspondence: amy.linsky@va.gov
[1]Section of General Internal Medicine, VA Boston Healthcare System, 150 S. Huntington Ave, Boston, MA 02130, USA
[2]Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Ave, Boston, MA 02118, USA
Full list of author information is available at the end of the article

at Veterans Affairs (VA) Boston Healthcare System from April 2010 until July 2011. Patients were former members of the United States military who sought and were eligible to receive care at the VA. Data were collected by fourth-year medical students, who were individually instructed on project processes as part of their Ambulatory Medicine Quality Improvement rotation. One student per month was assigned to the rotation. Immediately prior to a student-led clinical encounter, patients were given a printout of the electronic health record (EHR) listing of their medications and asked to answer two questions: "Which one of your medicines, if any, do you think is the most important? (if none, please write 'none')" and "Which one of your medicines, if any, do you think is the least important? (if none, please write 'none')." If needed, the student assisted the Veterans by reading the questions or writing their responses.

Because the data collection was intended as a quality improvement educational project, informed consent was not obtained. One of us (AL) entered all data into an Excel database, and we analyzed only the first chronologic encounter for each Veteran. Patient responses were entered exactly as written. Specific medications identified by patients were classified into medication classes using VA Drug Class Codes. Other data extracted from the EHR included patient sex, age at visit and number of actively prescribed medications.

Our two primary outcomes were patients' responses to the "most important" and "least important" questions. Responses to each question were categorized as one of four types: 1) One specific medication – the patient identified one medication only; 2) More than one medication – the patient reported more than one medication, chose medications to treat a diagnosis (e.g. "heart meds") or wrote "all"; 3) None – the patient wrote "none"; and 4) Did not answer the question – the patient left the response blank or wrote "n/a," "don't know," "uncertain" or "not sure."

We determined the prevalence of each primary outcome. Frequency counts identified the medication classes involved for responses where specific medications were chosen. Finally, we used chi-square tests to assess for associations between patient factors and choosing one specific medication as most or least important. All analyses were performed with SAS version 9.2 (SAS Institute, Inc) or Excel (Microsoft). Statistical significance was set at alpha < 0.05. This study protocol was approved by the Institutional Review Board of the VA Boston Healthcare System.

Results

Our study cohort of 104 Veterans was predominantly male (95%), and 59 (57%) were age 65 years or older (Table 1). Lists contained a mean of 9 medications (SD 5.7). Responding to the question about their most important medication, 41 patients (39%) identified one specific medication; 26 (25%) selected more than one; 21 (20%) wrote "none"; and 16 (15%) did not answer the question (Table 2). Answering the question about their least important medication, 31 Veterans (30%) chose one specific medication; two (2%) chose more than one; 51 (49%) wrote "none"; and 20 (19%) did not directly answer the question (Table 2).

There was no association between selecting one medication as most or least important and any of the available patient factors (data not shown). The most commonly cited most important medication classes were beta blockers (n = 8), angiotensin converting enzyme inhibitors (n = 7) and anticoagulants (n = 5). The most commonly cited least important medication classes were vitamins (n = 12), non-opioid analgesics (n = 4), antilipemic agents (n = 2) and antigout agents (n = 2). Full results of medications classes identified by patients as most important and least important are shown in Table 3.

Table 1 Baseline characteristics (n=104)

Patient characteristics	n (%)
Age 65 or older	59 (57)
Male	99 (95)
Total number of medications [mean (SD)]	9 (5.7)

*Care may have occurred previously or concurrently with present care at VA Boston.
†Medications dispensed from other VA facilities (i.e., remotely-dispensed), non-VA medications or documented as inpatient medications.

Table 2 Categories of responses to identification of most and least important medication

Response category	Most important n (%)	Least important n (%)
One specific medication	41 (39)	31 (30)
More than one medication*	26 (25)	2 (2)
More than one medication	5 (5)	0 (0)
Chose medications for a condition (did not name a specific medication)	11 (11)	1 (1)
More than one medication and chose it by condition	3 (3)	0 (0)
Wrote "all"	7 (7)	1 (1)
Wrote "None"	21 (20)	51 (49)
Did not answer the question*	16 (15)	20 (19)
Wrote "n/a"	4 (4)	5 (5)
Left it blank	10 (10)	11(11)
Wrote "don't know," "uncertain," or "not sure"	1 (1)	4 (4)
Wrote something undecipherable	1 (1)	0 (0)

*Subcategories sum to the total for their respective categories.

Table 3 Medication classes chosen as most or least important

Most important medication	n*
Beta blockers	8
Angiotensin converting enzyme inhibitors	7
Anticoagulants	5
Opioid analgesics	4
Calcium channel blockers	3
Non-opioid analgesics	2
Gastric medications	2
Digitalis glycosides	2
Antilipemic agents	2
Glucocorticoids	2
Insulin	2
Loop diuretics	2
Platelet aggregation inhibitors	1
Sedatives/hypnotics	1
Anticonvulsants	1
Antiparkinson agents	1
Antianginals	1
Thiazides/related diuretics	1
Local anesthetics, topical	1
Digestants	1
Genitourinary agents	1
Oral hypoglycemic agents	1
Antirheumatic agents	1
Skeletal muscle relaxants	1
Bronchodilators	1
Least important medication	**n**
Vitamins	12
Non-opioid analgesics	4
Antilipemic agents	2
Antigout agents	2
Sedatives/Hypnotics	1
Anticonvulsants	1
Loop diuretics	1
Angiotensin converting enzyme inhibitors	1
Local anesthetics, topical	1
Laxatives	1
Histamine antagonists	1
Antispasmodics, urological	1
Contraceptives, systemic	1
Skeletal muscle relaxants	1
Iron	1

*Unit of analysis is medication. Results do not include responses for condition-related medications (i.e., "heart meds").

Discussion

Approximately one in three Veteran patients did not identify a most important medication, and more than two in three patients did not identify a least important medication. There are several possible reasons to explain this finding. Patients may not have understood the questions, leading them to choose medications to treat a particular condition – therefore not selecting a single medication – or simply to leave the response blank. Limited health literacy, reflected as poor understanding of their health conditions or of their medications and associated indications, may have contributed to low response rates [8-10]. Without full understanding of their medical problems, the potential consequences of untreated health conditions and the possible benefits and risks of medication for those conditions, some patients may have been unable to make an informed selection from their medication list. This notion highlights the importance of education and knowledge in enabling patients to maintain an active role in their health care decisions.

Another hypothesis for the observed pattern of responses is that the use of the word "important" may not accurately reflect the construct that we were attempting to measure. That is, patients may have perceived medications as important but may not have believed that they are of a necessity or benefit to their future health to warrant full adherence. Others have demonstrated that perceived adverse effects of medications often outweigh preventive benefits of medications [11]. Competing demands, such as financial or social obligations, may further outweigh patient perceptions of importance. Psychometric testing of the best way to assess patient prioritization of medications with reliability and validity will improve measurement in future studies.

Another possible explanation for our findings is that Veteran patients were in fact able to understand and identify a medication yet they were unwilling to share these beliefs with their providers. Patients discontinue prescriptions for a variety of reasons without informing their healthcare provider of this decision [12]. Improving communication between patients and providers can lead to better shared decision making and better adherence [2].

Our study results need to be interpreted in the context of several limitations. We analyzed a small convenience sample of patients from one site within a larger health care system, and the study participants' responses may not reflect the views of Veterans receiving care elsewhere or the perceptions of non-Veterans. Future research involving multiple settings and patient populations will enable better generalizability. We also had limited information on patients' comorbidities and other medications, restricting us from appraising patient

responses as concordant with clinician opinion. However, these answers still reflect what the patient perceived, and it is recognized that patient beliefs are associated with medication adherence [4]. Additionally, explicit discussions may enable providers to better reconcile known conflicts between what they believe is clinically best and what the patient perceives as important [13,14]. Finally, the training status of providers (i.e., medical students) theoretically could have influenced patients' willingness to divulge their prioritization.

In this study population where the mean number of medications was nine, higher than most commonly accepted definitions of polypharmacy [15], a minority of patients were able to express a medication as least important. Better understanding of how patients prioritize their medications and how best to elicit this information will improve patient-provider communication and perhaps lead to discontinuation of medications that both the patient and the clinician feel have less importance.

Abbreviations

VA: Veterans affairs; EHR: Electronic health record (EHR); SD: Standard deviation.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AL was responsible for study concept and design; acquisition, analysis and interpretation of data; and drafting and critical revision of the manuscript. SRS was involved in analysis and interpretation of data and critical revision of the manuscript. AL had full access to all of the data in the study and takes responsibility for the integrity and accuracy of the data analysis. All authors read and approved the final manuscript.

Acknowledgements

We thank Anand Kartha, MD and Pamela Bellino, MA for the opportunity to engage medical students, as well as all medical students who participated. The study was conducted using resources of the VA Boston Healthcare System. The work was initiated while AL was a fellow supported by Boston Medical Center and the Center for Organization, Leadership, and Management Research at VA Boston Healthcare System. The funding organization had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or in the decision to submit the manuscript. Findings from this paper have been presented at the Society for General Internal Medicine annual meeting (Orlando, May 2012) and at the Academy Health annual research meeting (Orlando, June 2012).

Author details

[1]Section of General Internal Medicine, VA Boston Healthcare System, 150 S. Huntington Ave, Boston, MA 02130, USA. [2]Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Ave, Boston, MA 02118, USA. [3]Boston University School of Medicine, Boston, USA. [4]Division of General Internal Medicine, Brigham and Women's Hospital, Boston, USA. [5]Harvard Medical School, Boston, MA, USA.

Received: 5 September 2012 Accepted: 30 October 2012
Published: 2 November 2012

References

1. Kripalani S, Yao X, Haynes RB: **Interventions to enhance medication adherence in chronic medical conditions: a systematic review.** *Arch Intern Med* 2007, **167**(6):540–550.
2. Vermeire E, Hearnshaw H, Van Royen P, Denekens J: **Patient adherence to treatment: three decades of research. A comprehensive review.** *J Clin Pharm Ther* 2001, **26**(5):331–342.
3. Pladevall M, Brotons C, Gabriel R, Arnau A, Suarez C, de la Figuera M, Marquez E, Coca A, Sobrino J, Divine G, *et al*: **Multicenter cluster-randomized trial of a multifactorial intervention to improve antihypertensive medication adherence and blood pressure control among patients at high cardiovascular risk (the COM99 study).** *Circulation* 2010, **122**(12):1183–1191.
4. DiMatteo MR, Haskard KB, Williams SL: **Health beliefs, disease severity, and patient adherence: a meta-analysis.** *Medical care* 2007, **45**(6):521–528.
5. Fried TR, McGraw S, Agostini JV, Tinetti ME: **Views of older persons with multiple morbidities on competing outcomes and clinical decision-making.** *J Am Geriatr Soc* 2008, **56**(10):1839–1844.
6. Piette JD, Heisler M, Harand A, Juip M: **Beliefs about prescription medications among patients with diabetes: variation across racial groups and influences on cost-related medication underuse.** *J Health Care Poor Underserved* 2010, **21**(1):349–361.
7. Piette JD, Heisler M, Krein S, Kerr EA: **The role of patient-physician trust in moderating medication nonadherence due to cost pressures.** *Arch Intern Med* 2005, **165**(15):1749–1755.
8. Modig S, Kristensson J, Ekwall AK, Hallberg IR, Midlov P: **Frail elderly patients in primary care–their medication knowledge and beliefs about prescribed medicines.** *Eur J Clin Pharmacol* 2009, **65**(2):151–155.
9. Spiers MV, Kutzik DM, Lamar M: **Variation in medication understanding among the elderly.** *Am J Health Syst Pharm AJHP: official J Am Soc Health Syst Pharm* 2004, **61**(4):373–380.
10. Shrank WH, Avorn J: **Educating patients about their medications: the potential and limitations of written drug information.** *Health Aff* 2007, **26**(3):731–740.
11. Fried TR, Tinetti ME, Towle V, O'Leary JR, Iannone L: **Effects of benefits and harms on older persons' willingness to take medication for primary cardiovascular prevention.** *Arch Intern Med* 2011, **171**(10):923–928.
12. Moen J, Bohm A, Tillenius T, Antonov K, Nilsson JL, Ring L: **"I don't know how many of these [medicines] are necessary." - a focus group study among elderly users of multiple medicines.** *Patient Educ Couns* 2009, **74**(2):135–141.
13. Fried TR, Tinetti ME, Iannone L: **Primary care clinicians' experiences with treatment decision making for older persons with multiple conditions.** *Arch Intern Med* 2011, **171**(1):75–80.
14. Heisler M, Vijan S, Anderson RM, Ubel PA, Bernstein SJ, Hofer TP: **When do patients and their physicians agree on diabetes treatment goals and strategies, and what difference does it make?** *J Gen Intern Med* 2003, **18**(11):893–902.
15. Bushardt RL, Massey EB, Simpson TW, Ariail JC, Simpson KN: **Polypharmacy: misleading, but manageable.** *Clin Interv Aging* 2008, **3**(2):383–389.

doi:10.1186/1756-0500-5-619
Cite this article as: Linsky and Simon: Patients' perceptions of their "most" and "least" important medications: a retrospective cohort study. *BMC Research Notes* 2012 **5**:619.

Digestive and Liver Disease 44 (2012) 497–503

Contents lists available at SciVerse ScienceDirect

Digestive and Liver Disease

journal homepage: www.elsevier.com/locate/dld

Liver, Pancreas and Biliary Tract

Primary care-based interventions are associated with increases in hepatitis C virus testing for patients at risk☆

Alain H. Litwin[a,b,*], Bryce D. Smith[e], Mari-Lynn Drainoni[f,g,i], Diane McKee[d], Allen L. Gifford[f,h,j,i], Elisa Koppelman[f,i], Cindy L. Christiansen[f,i], Cindy M. Weinbaum[e], William N. Southern[a,b,c]

[a] Department of Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, United States
[b] Division of General Internal Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, United States
[c] Division of Hospital Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, United States
[d] Department of Family Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, United States
[e] Division of Viral Hepatitis, Centers for Disease Control and Prevention, National Center for HIV/Viral Hepatitis/STD/TB Prevention, United States
[f] Health Policy and Management Department, School of Public Health, Boston University, United States
[g] Section of Infectious Diseases, Boston University School of Medicine, United States
[h] Section of General Internal Medicine, Boston University School of Medicine, United States
[i] Center for Health Quality, Outcomes and Economic Research, ENRM Veterans Administration Hospital, United States
[j] VA QUERI-HIV/Hepatitis Program, Edith Nourse Rogers Memorial Veterans Hospital, United States

ARTICLE INFO

Article history:
Received 24 August 2011
Accepted 16 December 2011
Available online 18 February 2012

Keywords:
Chronic hepatitis C
HCV
Primary care
Public health
Screening

ABSTRACT

Background: An estimated 3.2 million persons are chronically infected with the hepatitis C virus (HCV) in the U.S. Effective treatment is available, but approximately 50% of patients are not aware that they are infected. Optimal testing strategies have not been described.
Methods: The Hepatitis C Assessment and Testing Project (HepCAT) was a serial cross-sectional evaluation of two community-based interventions designed to increase HCV testing in urban primary care clinics in comparison with a baseline period. The first intervention (risk-based screener) prompted physicians to order HCV tests based on the presence of HCV-related risks. The second intervention (birth cohort) prompted physicians to order HCV tests on all patients born within a high-prevalence birth cohort (1945–1964). The study was conducted at three primary care clinics in the Bronx, New York.
Results: Both interventions were associated with an increased proportion of patients tested for HCV from 6.0% at baseline to 13.1% during the risk-based screener period ($P < 0.001$) and 9.9% during the birth cohort period ($P < 0.001$).
Conclusions: Two simple clinical reminder interventions were associated with significantly increased HCV testing rates. Our findings suggest that HCV screening programs, using either a risk-based or birth cohort strategy, should be adopted in primary care settings so that HCV-infected patients may benefit from antiviral treatment.

1. Introduction

An estimated 3.2 million persons are chronically infected with the hepatitis C virus (HCV) in the U.S. Although the prevalence of HCV infection is estimated at 1.6% in the U.S. [1], urban clinic populations may have an HCV prevalence as high as 8% [2,3]. HCV infection causes approximately 40% of chronic liver disease [4] and the majority of cases of hepatocellular carcinoma [5], and we can expect that there will be up to a 4-fold increase in HCV-related deaths over the next 20 years [6–8]. Effective treatment is available for HCV which can lead to long-term remission of disease and decreased liver-related mortality rates, but approximately 50% of patients are not aware that they are infected [9,10]. Although the Institute of Medicine (IOM) has recommended increased HCV testing [11], optimal testing strategies have not been described [12].

☆ The project was funded by the Agency for Health Care Research and Quality (AHRQ) and the Centers for Disease Control and Prevention via the AHRQ ACTION initiative to Boston University, contract HHSA2902006000012 T0#4. The content is solely the responsibility of the authors and does not necessarily represent the official views of AHRQ. The study and investigators were also funded by the Clinical Investigation Core of the Center for AIDS Research at the Albert Einstein College of Medicine and Montefiore Medical Center, funded by the National Institutes of Health NIH P30 AI51519, NIDA K23 DA022454, and the CTSA Grant UL1 RR025750 and KL2 RR025749 and TL1 RR025748 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH).
* Corresponding author at: Department of Medicine, Montefiore Medical Center, 111 East 210th Street, Bronx, NY 10467, USA. Tel.: +1 718 944 3862.
E-mail address: alitwin@montefiore.org (A.H. Litwin).

1590-8658/$36.00 © 2011 Editrice Gastroenterologica Italiana S.r.l. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.dld.2011.12.014

The aim of this analysis is to evaluate two interventions designed to increase HCV testing for patients at risk for HCV infection and disease.

There are several potential screening strategies which may be used when designing HCV testing programs. The Centers for Disease Control and Prevention and other organizations have recommended HCV testing for persons with an increased risk of infection such as persons who inject drugs [4,13,14]. Risk-based approaches to testing may be efficient, because patients with identified risks such as injection drug use have a much higher prevalence of HCV infection (60–90%) [15] than the general population (1.6%) [1]. However, risk-based screening has been limited because primary care providers often do not ask patients about HCV risk factors, testing is rarely initiated based on physician-identified risks, and patients are often reticent to disclose stigmatized risk behaviors [16,17].

Another approach to HCV screening is to identify a group of patients at higher risk without asking about stigmatized risk behaviors. Americans born between 1945–1964 are more likely to be HCV-infected than those in other age groups [1], and we have established at the three study clinics that HCV infection is nearly 4 times more likely in this high-prevalence birth cohort compared to patients outside of this birth cohort [3]. Approximately 76% of HCV-infected Americans were born between 1945–1964, and Americans born between 1945 and 1964 have a anti-HCV prevalence as high as 4.3% [1]. Therefore, another strategy is to test all patients born between 1945 and 1964.

The Hepatitis C Assessment and Testing Project (HepCAT) was a prospective evaluation funded by CDC via AHRQ ACTION of two serial community-based interventions designed to increase rates of HCV testing of at-risk or high-prevalence patients in urban primary care clinics from rates seen during baseline testing (baseline period). The first intervention (risk-based screener intervention) prompted physicians with a clinical reminder sticker in the progress note to ask whether a patient had specific HCV-related risks and to order HCV tests based on the presence of these risks. The second intervention (birth cohort intervention) prompted physicians with a clinical reminder sticker in the progress note to order HCV tests on all patients born within a high-prevalence birth cohort (1945–1964). We hypothesized that both interventions would be associated with an increase in testing of patients for HCV.

2. Methods

2.1. Study design

The study was a serial cross-sectional evaluation of two community-based interventions. We evaluated HCV testing and positivity rates during a baseline and two subsequent intervention periods. We used electronic medical record (EMR) data to examine the associations of interventions with HCV testing rates overall and within at-risk sub-groups.

2.2. Study setting

The study was conducted at three community-based primary care (family medicine or internal medicine) clinics affiliated with Montefiore Medical Center, a university-affiliated teaching hospital. The three participating primary care clinics are large, urban clinics located in the Bronx, New York. Each year, 54,000 adults attend over 150,000 primary care visits to the three clinics. Physicians were 153 internal medicine and family medicine physicians, both attendings and residents. The clinic sites are located in economically depressed areas of the Bronx and serve patients with high rates of poverty and substance use. Reported prevalence of HCV infection is higher in New York City than the national estimate and the Bronx has a higher prevalence than NYC as a whole [18,19].

2.3. Study population

We examined data on all adult patients (at least 18 years of age) who had a primary care visit to one of three clinics during at least one of the three study periods. The three study periods included the baseline period (January 1, 2008–February 28, 2008), risk-based screener intervention period (November 24, 2008–March 6, 2008), and birth cohort intervention period (March 9, 2009–June 30, 2009). We excluded patients who had been previously tested for HCV at Montefiore Medical Center.

2.4. Description of the interventions

2.4.1. Components common to both interventions

Both the risk-based screener and birth cohort interventions included: (1) on-site educational sessions for primary care providers (PCPs) and staff – delivered prior to and during interventions; (2) regular communication between the research team and clinical leadership (by phone or email); and (3) environmental reminders (HepCAT buttons, pocket cards, and posters). In addition, project staff visited each clinic twice-weekly to place stickers on all progress notes, encourage adherence to screening protocols, and elicit feedback from clinic staff.

2.4.2. Risk-based screener intervention

The screener intervention included a risk-based screener sticker which was placed on top of each progress note (Supplementary Fig. S1). The screener prompted the physician to ask the patient nine questions related to HCV risk: behavioral risk factors (ever injected or snorted any drug); risk group associations (ever incarcerated or ever homeless); medical risk factors (transfusion or organ transplant before 1992, told by physician that patient had liver disease, long-term hemodialysis), and other risk factors (maternal hepatitis C). In addition, the physician was asked to indicate whether patients ever had an elevated ALT (defined as female $\geq$ 20 IU/L or male $\geq$ 31 IU/L) [20]. These risks were chosen based on CDC HCV testing recommendations, American Association for the Study of Liver Diseases (AASLD) guidelines and other associations reported in the literature [13,14,21]. Physicians were asked to complete the sticker at every patient visit unless previously completed, and to order an HCV antibody test if any risk was identified. Patients who were already HCV antibody positive or who had been tested within the last 12 months were not asked to be tested. Spanish translations of the sticker were available in every medical office, and laminated versions were placed in each provider's mailbox. PCPs were supplied with a script (English and Spanish) to help standardize and normalize the introduction of the screening questions. This phase was conducted from November 24, 2008 to March 6, 2009 (15 weeks).

2.4.3. Birth cohort intervention

During the birth cohort intervention, a birth cohort reminder sticker was placed on top of each progress note (Supplementary Fig. S2). The birth cohort reminder prompted PCPs to order HCV antibody tests for all patients born between 1945 and 1964 regardless of any other identified risk. The risk-based screener sticker was no longer placed on top of progress notes. Spanish translations of the birth cohort sticker and provider scripts (in both English and Spanish) were available in every provider office. This phase was conducted from March 9, 2009 to June 30, 2009 (16 weeks).

Table 1
Characteristics of study populations.

	Baseline ($n = 6591$)	Risk-based ($n = 8981$)	Birth-cohort ($n = 10{,}165$)	P-value
Age	48.3 ± 17.7	47.8 ± 17.7	47.4 ± 17.8	0.28
Male	1622 (24.6)	2330 (25.9)	2636 (25.9)	0.10
Race/ethnicity				0.06
White	330 (5.0)	389 (4.3)	442 (4.3)	
Black	2064 (31.3)	2733 (30.4)	3039 (30.0)	
Latino	3349 (50.8)	4734 (52.7)	5395 (53.0)	
Oth/unknown	848 (12.9)	1125 (12.5)	1289 (12.7)	
Insurance[a]				<0.001
Medicare	757 (11.5)	1029 (11.5)	1140 (11.2)	
Medicaid	3491 (53.0)	4609 (51.3)	5211 (51.3)	
Commercial	1596 (24.2)	2062 (23.0)	2404 (23.6)	
Self	740 (11.2)	1272 (14.2)	1391 (13.7)	
Diagnoses				
Substance abuse[b]	219 (3.3)	292 (3.3)	342 (3.4)	0.91
Alcohol abuse[c]	72 (1.1)	101 (1.1)	103 (1.0)	0.74
HIV[d]	148 (2.2)	167 (1.9)	174 (1.7)	0.04
STD[e]	154 (2.3)	244 (2.7)	303 (3.0)	0.04
Cirrhosis[f]	25 (0.4)	31 (0.3)	38 (0.4)	0.92
ESRD[g]	17 (0.3)	28 (0.3)	27 (0.3)	0.77
Psychiatric diagnosis[h]	918 (13.9)	1355 (15.1)	1608 (15.8)	0.004

Continuous variables reported as mean ± standard deviation, tested using Kruskal–Wallis test dichotomous variables reported as No. (%), tested using Chi-squared test.
[a] Column does not add to 100% because of missing values.
[b] ICD-9 or positive urine toxicology.
[c] ICD-9 for Etoh dependance or etoh liver disease or etoh level ≥ 80.
[d] ICD-9 or positive antibody test or Western blot.
[e] STD = sexually transmitted disease (not HIV): ICD-9 or positive GC or chlamydia PCR probe.
[f] ICD-9 code.
[g] ESR = end-stage renal disease: ICD-9 code or procedure code for hemodialysis.
[h] ICD-9 for affective, anxiety, schizophrenia, or psychosis.

2.5. Data extraction

We extracted demographic and clinical information dating back to March 1997, the year electronic records became available, including inpatient and outpatient ICD-9 diagnosis codes and laboratory testing results. The Institutional Review Boards of Boston University Medical Center and Montefiore Medical Center approved this study.

2.6. Outcome variables

In these analyses, we included only patients not previously tested for HCV. Patients were considered not to be previously tested for HCV if there was no HCV antibody test result in the EMR dating back to March 1997. The primary outcome was HCV antibody testing during each intervention period, and we also examined rates of HCV antibody positivity. HCV antibody testing was defined as an anti-hepatitis C virus antibody (anti-HCV) by ELISA performed either on the same date of the primary care visit or within 3 months of the initial primary care visit within the baseline or intervention periods. HCV antibody positivity (indicating past or current HCV infection) was defined as a positive anti-HCV test performed during that specified time period.

2.7. Other variables

We included the following sociodemographic variables. Age was dichotomized as within the high prevalence birth cohort (born between 1945 and 1964) vs. not within the cohort. Insurance status was categorized based on primary insurance: Medicare, Medicaid, Commercial, and Self/Uninsured. Race/ethnicity was collapsed into four categories: non-Hispanic White, non-Hispanic Black or African American, Latino or Hispanic, and other/unknown.

For at-risk sub-group analysis, we examined the variables below which were defined with ICD-9 codes and/or laboratory values. ICD-9 codes were classified using the Healthcare Cost and Utilisation Project of the Agency for Healthcare Research and Quality system [22]. Although a history of blood transfusion or organ transplant before 1992 is a known risk factor for HCV infection, the EMR did not have data on these risks, so the analysis does not include them. The following conditions were considered present if recorded at any time from March 1997 through the qualifying visit date.

Substance abuse. ICD-9 code for substance abuse/dependence or positive urine toxicology for amphetamines, barbiturates, cocaine, or methadone.

HIV. ICD-9 code for HIV infection or a positive antibody test confirmed by Western blot.

Sexually transmitted disease (STD). ICD-9 code indicating gonorrhea or chlamydia or positive gonorrhea or chlamydia PCR probe.

Alcohol abuse. ICD-9 code for alcohol abuse/dependence or alcohol-related liver disease, or a serum alcohol level ≥ 80 mg/dL.

HBV infection: positive serological test for anti-HBc.

Cirrhosis. ICD-9 code for cirrhosis.

End stage renal disease (ESRD). ICD-9 code for end-stage renal disease or procedure code for hemodialysis.

Psychiatric disease. ICD-9 code for affective disorder, anxiety disorder, schizophrenia, or psychosis.

ALT elevation. The highest ALT value reported from March 1997 through the clinic visit date for each subject was used. Elevated ALT was treated as a dichotomous variable and defined in two different ways: (1) >40 U/L (40 U/L is a commonly used upper limit of normal for all adults) [23] and (2) >19 U/L for females and >30 U/L for males (updated upper limits of normal) [20].

Any HCV risk. Any HCV risk was defined as having at least one of the following risks as defined above: substance abuse; HIV; STD; cirrhosis; ESRD, or elevated ALT (>40).

2.8. Statistical analysis

2.8.1. Proportion tested

Chi square and ANOVA tests were used to examine differences in demographic and clinical characteristics among the baseline period, risk-based screener period, and birth cohort period.

In each of the three periods, we determined the proportion of patients tested for anti-HCV. The proportions tested were calculated for predefined age categories and demographic characteristics, presence or absence of at-risk sub-groups, and the presence or absence of ALT elevation.

2.8.2. Comparison of intervention HCV testing rates with baseline testing rates

We compared the rates of HCV testing in the risk-based screener and birth cohort intervention periods with the baseline testing period, overall and within multiple subgroups by using chi square tests. We examined the odds of HCV testing for the risk-based screener and birth cohort intervention periods as compared with the baseline period, overall and within multiple subgroups using simple logistic regression models. To compare odds of HCV testing in each intervention period with the baseline period, we constructed multivariate models to adjust for age, sex, race/ethnicity, insurance status, substance use, alcohol abuse, HIV, STDs, cirrhosis, psychiatric diagnosis, and elevated ALT. To assess the influence of cross-contamination across periods, we performed multiple sensitivity analyses excluding patients who appeared in more than one period, and including only patients seen in more than one period, as well as all three periods.

STATA/IC software, version 10.0 (StataCorp, College Station, TX) was used for all data management and statistical analysis.

3. Results

3.1. Study population

The baseline cohort ($n = 6591$), the risk screener cohort ($n = 8981$), and the birth cohort ($n = 10{,}165$) included all adult patients (at least 18 years of age) who had not been tested for anti-HCV in the past (since 1997), and who made at least one primary care visit during the baseline, risk screener intervention, or birth cohort intervention periods respectively.

Demographic and clinical information for the overall study population are summarized in Table 1. The mean age was 47.8 years (SD = 17.7). The population was primarily female (74.4%), Latino (52.4%) or African American (30.4%), and Medicaid-insured (51.7%). Overall, 15.1% had a history of psychiatric disease, 3.3% had a history of substance abuse, and 1.9% had a history of HIV. The baseline period, risk-based screener period, and birth cohort period were similar with regard to most demographics and risks; slight differences were observed in rates of Latino patients, insurance status, HIV, STDs, and patients with psychiatric diagnoses.

3.2. HCV testing rates in each intervention period

Both intervention periods (Table 2) were associated with an increased proportion of patients tested for HCV from 6.0% at baseline to 13.1% in the risk screener phase ($P < 0.001$) and 9.9% in the birth cohort phase ($P < 0.001$). After adjustment, both interventions were associated with significantly increased odds of testing for HCV: aOR 2.37 (95% CI 2.10–2.67) for the risk-based screener period and aOR 1.70 (95% CI 1.50–1.92) for the birth-cohort period. To assess the influence of cross-contamination across periods, we performed sensitivity analyses excluding patients who appeared in more than one phase, and including only patients seen in more than one phase, as well as all three phases. In each sensitivity analysis, the results were similar to the primary analysis.

When the analysis was stratified by clinic, each clinic experienced a significant increase in the rate of HCV testing. In the clinic with the lowest baseline rate of testing, HCV testing increased almost four fold during the risk-based screener period (2.9% versus 11.3%, $P < 0.001$) and three fold during the birth cohort period (2.9% versus 8.7%, $P < 0.001$).

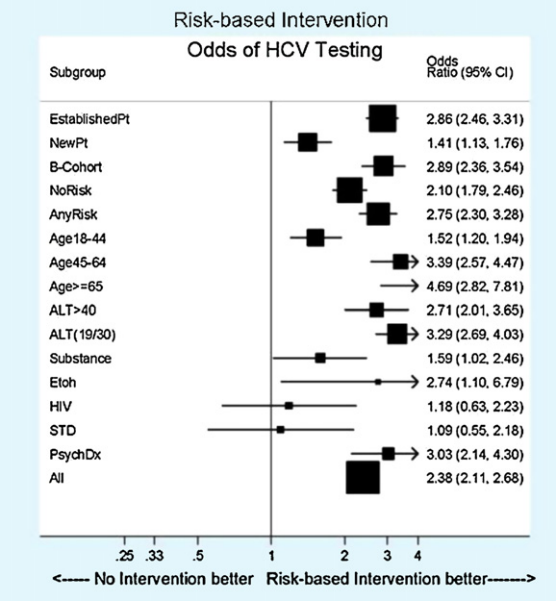


Fig. 1. Risk-based screener versus baseline forest plot.

3.3. Sub-group analysis

3.3.1. Risk-based screener intervention

Rates of anti-HCV screening increased in the following subgroups: patients with at least one risk (5.0–12.7%, $P < 0.001$), elevated ALT (5.7–14.1%, $P < 0.001$), HBV infection (1.6–8.5%, $P < 0.001$), 1945–1964 birth cohort (5.4–14.3%, $P < 0.001$), substance abuse or dependence (17.3–25.0%, $P = 0.04$), alcohol abuse or dependence (9.7–22.8%, $P = 0.02$), and psychiatric diagnosis (4.6–12.7%, $P < 0.001$). Rates of HCV testing significantly increased in all racial/ethnic groups and insurance sub-groups. There were significant increased adjusted odds of HCV testing in almost all sub-groups (Fig. 1).

3.3.2. Birth cohort intervention

Rates of anti-HCV screening increased in the following subgroups: patients with at least one risk (5.0–7.8%, $P < 0.001$), elevated ALT (5.7–8.1%, $P = 0.02$), previous HBV infection or vaccination (1.6–5.2%, $P = 0.02$), 1945–1964 birth cohort (5.4–13.2%, $P < 0.001$), and psychiatric diagnosis (4.6–8.7%, $P < 0.001$). Rates of HCV testing significantly increased in all racial and ethnic groups except Whites and in all insurance groups except those who were self-insured. There were significant increased adjusted odds of HCV testing in almost all sub-groups (Fig. 2).

3.4. Yield of HCV testing in each phase

Both interventions were associated with rates of anti-HCV-positivity that were more than three times the national prevalence of 1.6% [1]. The rate of anti-HCV-positivity was 5.3% in the risk-based screener period and 5.8% in the birth cohort period. Overall, 36 new anti-HCV+ patients (0.5% of all previously untested patients) were identified during the baseline period, 62 (0.7% of all previously untested patients) during the risk-based screener intervention period, and 59 (0.6% of all previously untested patients) during the birth cohort period, but the differences were not statistically significant.

Table 2
Proportion of patients tested.

	Baseline (n = 6591)	Risk-based (n = 8981)	Birth-cohort (n = 10,165)	Base/risk P-value	Base/birth P-value
All patients never tested – No.(%)	394 (6.0)	1179 (13.1)	1008 (9.9)	<0.001	<0.001
Clinic A (n = 4786)	5.6	14.0	8.7	<0.001	<0.001
Clinic B (n = 5032)	10.4	14.7	12.7	<0.001	0.01
Clinic C (n = 6938)	2.9	11.3	8.7	<0.001	<0.001
Established patients (n = 14,852)	4.1	10.8	7.6	<0.001	<0.001
New patients (n = 1904)	22.3	28.9	26.8	0.002	0.03
Within birth cohort (n = 5910)	5.4	14.3	13.2	<0.001	<0.001
Not within birth cohort (n = 10,846)	6.3	12.5	8.2	<0.001	<0.001
No risk factor (n = 7726)	7.0	13.5	12.0	<0.001	<0.001
Any risk factor[a] (n = 7846)	5.0	12.7	7.8	<0.001	<0.001
Age 18–44	8.5	14.5	9.9	<0.001	0.04
Age 45–64	5.2	14.3	13.2	<0.001	<0.001
Age ≥ 65	1.9	7.8	4.1	<0.001	<0.001
Females	4.7	10.7	8.0	<0.001	<0.001
Males	10.0	20.1	15.5	<0.001	<0.001
Race/ethnicity					
White	3.9	11.6	5.0	<0.001	0.49
Black	6.4	13.5	9.2	<0.001	<0.001
Latino	6.2	12.7	10.5	<0.001	<0.001
Oth/unknown (n = 1973)	4.9	14.5	11.0	<0.001	<0.001
Insurance					
Medicare	2.0	9.2	4.6	<0.001	0.003
Medicaid	5.5	12.9	9.8	<0.001	<0.001
Commercial	6.3	13.4	10.9	<0.001	<0.001
Self	11.6	16.7	13.1	0.002	0.33
ALT > 40	5.7	14.1	8.1	<0.001	0.02
ALT elevation (19/30)	3.9	11.9	6.9	<0.001	<0.001
Anti-HBc	1.6	8.5	5.2	<0.001	0.02
Substance abuse[b]	17.3	25.0	17.8	0.04	0.88
Alcohol abuse[c]	9.7	22.8	8.7	0.02	0.82
HIV[d]	13.5	15.6	13.2	0.61	0.94
STD[e]	9.1	9.8	6.3	0.80	0.27
Cirrhosis[f]	4.0	9.7	13.2	0.41	0.23
ESRD[g]	0.0	7.1	18.5	0.26	0.06
Psychiatric diagnosis[h]	4.6	12.7	8.7	<0.001	<0.001

[a] Substance abuse, HIV, STD, cirrhosis, ESRD, or elevated ALT(19/30).
[b] ICD-9 or positive urine toxicology.
[c] ICD-9 for Etoh dependance or etoh liver disease or etoh level ≥ 80.
[d] ICD-9 or positive antibody test or Western blot.
[e] STD = sexually transmitted disease (not HIV): ICD-9 or positive GC or chlamydia PCR probe.
[f] ICD-9 code.
[g] ESR = end-stage renal disease: ICD-9 code or procedure code for hemodialysis.
[h] ICD-9 for affective, anxiety, schizophrenia, or psychosis.

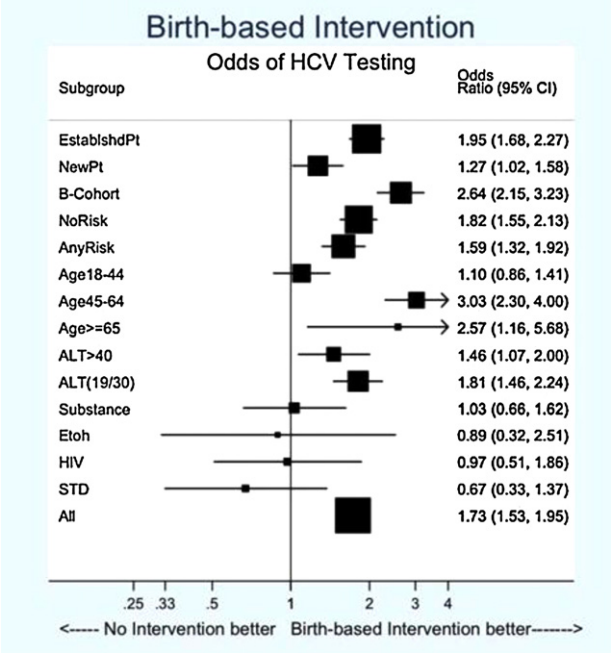


Fig. 2. Birth-cohort screener versus baseline forest plot.

4. Discussion

We demonstrated that implementation of a clinical reminder sticker as the backbone of a multi-component intervention was associated with a significant increase of anti-HCV testing for a population of at-risk and high prevalence patients in three large urban primary care clinics. Both the risk-based screener and birth cohort interventions were associated with significantly increased rates of HCV testing overall, and among those with HCV-related risks. In the clinic with the lowest rate of baseline testing, there was a four-fold increase in rates of HCV testing with the risk-based screener strategy and a three-fold increase in testing with the birth cohort strategy. Both interventions were associated with a considerable yield of testing (over 5% HCV-positivity).

Several recent studies have demonstrated that risk-based HCV testing can identify most HCV-positive patients through asking about specific risks, but only one study, by Zuniga et al., was integrated within routine patient care (in Veterans Health Administration setting where HCV testing was mandated by a change in federal policy), and none were designed to demonstrate an increase of HCV testing in high-risk patients compared with baseline testing practises [1,2,24]. There are several studies which demonstrate interventions that increase HIV testing in real-world settings. An integrated package of quality improvement interventions utilizing decision support (a real-time, electronic clinical reminder to

identify patients at increased risk of HIV infection), academic detailing and audit feedback resulted in a doubling of HIV testing for a population of at-risk individuals who had not been previously tested at two large Veterans Health Administration health care systems [25]. Similarly, our risk-based screener intervention led to a more than doubling of anti-HCV testing in individuals with identified risks who had not been previously tested.

Identification of patients who have been infected with hepatitis C is the first step toward increasing the number of patients who initiate HCV treatment. Patients who achieve a sustained viral response (SVR) with therapy will have long-term remission of disease, and liver-related mortality rates comparable to the general population [26–28]. In patients with HCV-related cirrhosis, regression of cirrhosis may occur in patients who are treated, and regression is associated with decreased disease-related morbidity and improved survival [27]. Identification of HCV-infected patients is urgently needed as the prevalence of hepatitis C cirrhosis and hepatocellular carcinoma has increased over the past ten years, and will continue to increase through the next decade [28,29]. At the current low rates of treatment, only 14.5% of the projected 259,000 HCV-related deaths between 2002 and 2030 will be prevented [9,30–32]. Based on the estimates of Volk et al., if treatment rates were increased to 75% and new treatment options including direct-acting antiviral medications achieve SVR 75% of the time, we can expect that over half of liver-related deaths could be avoided [9,33–36].

Our study had several strengths. First, by utilizing the EMR to define subgroups of patients with HCV-related risks, we were able to test the associated outcomes of a clinic-wide intervention on the entire patient population visiting the clinic during a specified period of time. Second, we found that incorporating these interventions was effective even in clinics where physicians already demonstrated high rates of testing at baseline [3]. Lastly, our study may be generalizable to many other real-world settings in that our interventions did not require extensive technology and were not integrated within the electronic medical record [37].

There were several limitations to our study. First, because the overall duration of both interventions was only 31 weeks, it is unclear if either intervention is sustainable. Next, because we did not utilize a contemporaneous comparison group, we were not able to establish a causal link between the interventions and the increased HCV testing observed. Our findings may not be generalizable to clinical settings outside of urban settings such as suburban private practice environments. Lastly, the study was powered to evaluate for change in HCV testing, not HCV testing yield, so we were unable to examine for differences in yield between the study periods.

Future directions

Future studies could investigate the effect of combining our strategies by adding the birth cohort question to the risk-based screener. Risk-based screening will allow identification of younger patients who are less likely to have comorbidities and contraindications to HCV treatment, and have a better chance of SVR. Conversely, birth cohort screening will allow identification of older patients who are more likely to urgently need antiviral treatment to prevent decompensated cirrhosis, hepatocellular carcinoma and death. Future studies should focus on the sustainability of HCV testing strategies, as adherence to our risk-based screener decreased over time, and risk-based screening may be less sustainable in a busy primary care practice than a birth cohort strategy [38]. The national strategy for HIV testing moved from risk-based to universal testing to increase timely identification of HIV-infected patients, and to reduce stigma [39,40]. Similarly, a universal testing strategy for HCV testing involving a birth cohort strategy will likely increase timely identification of HCV-infected patients and reduce stigma. We must also investigate the potential important downstream benefits of HCV testing programs (vaccinations; reduction in alcohol use; antiviral treatment; and reductions in HCV-related morbidity and mortality) [9,41,42].

In summary, we have found that two simple clinical reminder interventions were associated with significantly increased HCV testing rates when compared with pre-intervention testing rates. Identification is the first step to facilitate treatment of chronic hepatitis C, and possibly prevent HCV-related morbidity and mortality. To respond to the Institute of Medicine's report calling for increasing HCV testing in high-risk populations [11], HCV screening programs using either a risk-based or birth cohort strategy should become integrated within primary care settings so that we may realize the potentially life-saving benefits of treatment, and move actively toward eradication of HCV infection in the United States.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.dld.2011.12.014.

References

[1] Armstrong GL, Wasley A, Simard EP, et al. The prevalence of hepatitis C virus infection in the United States, 1999 through 2002. Ann Intern Med 2006;177:705–14.
[2] McGinn T, O'Connor-Moore N, Alfandre D, et al. Validation of a hepatitis C screening tool in primary care. Arch Intern Med 2008;168:2009–13.
[3] Southern WN, Drainoni ML, Smith BD, et al. Hepatitis C testing practices and prevalence in a high-risk urban ambulatory care setting. J Viral Hepatitis 2010, doi:10.1111/j. 1365-2893.2010.01327.
[4] Centers for Disease Control and Prevention. Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. Morb Mortal Wkly Rep 1998;47:1–33.
[5] El-Serag HB, Davila JA, Petersen NJ, et al. The continuing increase in the incidence of hepatocellular carcinoma in the United States: an update. Ann Intern Med 2003;139:817–23.
[6] Davis GL, Albright JE, Cook SF, et al. Projecting future complications of chronic hepatitis C in the United States. Liver Transpl 2003;9:331–8.
[7] Rein DB, Wittenborn JS, Weinbaum CM, et al. Forecasting the morbidity and mortality associated with prevalent cases of pre-cirrhotic chronic hepatitis C in the United States. Dig Liver Dis 2011;43:66–72.
[8] Wong JB, McQuillan GM, McHutchison JG, et al. Estimating future hepatitis C morbidity, mortality, and costs in the United States. Am J Public Health 2000;90:1562–9.
[9] Volk ML, Tocco R, Saini S, et al. Public health impact of antiviral therapy for hepatitis C in the United States. Hepatology 2009;50:1750–5.
[10] Wasley A, Finelli L, Bell B, et al. The knowledge and behavior of HCV-infected persons identified in a national seroprevalence survey, USA, 2001–2002. J Clin Virol 2006;36:S198–9.
[11] IOM (Institute of Medicine). Hepatitis and liver cancer: a national strategy for prevention and control of hepatitis B and C. Washington, DC: The National Academies Press; 2010.
[12] Lapane KL, Jakiche AF, Sugano D, et al. Hepatitis C infection risk analysis: who should be screened? Comparison of multiple screening strategies based on the National Hepatitis Surveillance Program. Am J Gastroenterol 1998;93:591–6.
[13] Dienstag JL, McHutchinson JG. American Gastroenterological Association medical position statement on the management of hepatitis C. Gastroenterology 2006;130:225–30.
[14] Ghany MG, Strader DB, Thomas DL, et al. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 2009;49:1335–74.
[15] Murrill CS, Weeks H, Castrucci BC, et al. Age-specific seroprevalence of HIV, hepatitis B virus, and hepatitis C virus infection among injection drug users admitted to drug treatment in 6 US cities. Am J Public Health 2002;92:385–7.
[16] Shehab TM, Sonnad SS, Lok ASF. Management of hepatitis C patients by primary care physicians in the United States: results of a national survey. J Viral Hepatitis 2001;8:377–83.
[17] Gunn RA, Murray PJ, Brennan CH, et al. Evaluation of screening criteria to identify persons with hepatitis C virus infection among sexually transmitted disease clinic clients: results from the San Diego Viral Hepatitis Integration Project. Sex Transm Dis 2003;30:340–4.
[18] Bornschlegel K, Berger M, Garg RK, et al. Prevalence of hepatitis C infection in New York city, 2004. J Urban Health 2009;86:909–17.

[19] Hepatitis A, B, and C Surveillance Report. New York City Department of Health and Mental Hygiene 2005. Accessed 24 September 2010 at: http://www.nyc.gov/html/doh/downloads/pdf/cd/cd-hepabc-surveillance-report.pdf.
[20] Prati D, Taioli E, Zanella A, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. Ann Intern Med 2002;137:1–10.
[21] Screening veterans for HCV infection. VA National Clinical Public Health Programs–Hepatitis C. Accessed 24 September 2010 at: http://www.hepatitis.va.gov/vahep?page=prtop02-cs-02.
[22] Healthcare Cost and Utilization Project, agency for Healthcare Research and Quality RMD. Clinical Classification Software (CCS), 2007. Available from: http://www.hcup-us.ahrq.gov/tools_software.jsp (Accessed February 5, 2012).
[23] Pratt DS, Kaplan MM. Evaluation of abnormal liver-enzyme results in asymptomatic patients. N Engl J Med 2000;342:1266–71.
[24] Zuniga IA, Chen CC, Lane DS, et al. Analysis of a hepatitis C screening programme for US veterans. Epidemiol Infect 2006;134:249–57.
[25] Goetz MB, Hoang T, Bowman C, et al. A system-wide intervention to improve HIV testing in the Veterans Health Administration. J Gen Intern Med 2008;23:1200–7.
[26] Kasahara A, Tanaka H, Okanoue T, et al. Interferon treatment improves survival in chronic hepatitis C patients showing biochemical as virological responses by preventing liver related death. J Viral Hepat 2004;11:148–56.
[27] Veldt BJ, Heathcote EJ, Wedemeyer H, et al. Sustained virologic response and clinical outcomes in patients with chronic hepatitis C and advanced fibrosis. Ann Intern Med 2007;147:677–84.
[28] Kanwal F, Hoang T, Kramer JR, et al. Increasing prevalence of HCC and cirrhosis in patients with chronic hepatitis C virus infection. Gastroenterology 2011;140:1182–8.
[29] Davis GL, Alter MJ, El-Serag H, et al. Aging of hepatitis C virus (HCV)-infected persons in the United States: a multiple cohort model of HCV prevalence and disease progression. Gastroenterology 2010;138:513–21.
[30] Yoshida H, Arakawa Y, Sata M, et al. Interferon therapy prolonged life expectancy among chronic hepatitis C patients. Gastroenterology 2002;123:483–91.
[31] Manns MP, McHutchison JG, Gordon SC, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomized trial. Lancet 2001;358:958–65.
[32] Fried MW, Shiffman ML, Reddy KR, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 2002;347:975–82.
[33] McHutchison JG, Everson GT, Gordon SC, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. N Engl J Med 2009;360:1827–38.
[34] Kwo PY, Lawitz EJ, McCone J, et al. Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naïve patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre phase 2 trial. Lancet 2010;376:705–16.
[35] Jacobson IM, McHutchison JG, Dusheiko G, et al. Telaprevir for previously untreated chronic hepatitis C infection. N Engl J Med 2011;364:2405–16.
[36] Poordad F, McCone J, Bacon BR, et al. Boceprevir for untreated chronic HCV genotype 1 infection. N Engl J Med 2011;364:1195–206.
[37] Burt CW, Sisk JE. Which physicians and practices are using electronic medical records? Health Aff 2005;24:1334–43.
[38] Southern WN, Drainoni ML, Smith BD, et al. Physician non-adherence with a hepatitis C screening program. In: Poster presented at the 62nd annual meeting for the american association for the study of liver diseases. 2011.
[39] Centers for Disease Control and Prevention. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. Morb Mortal Wkly Rep 2006;55:1–17.
[40] Qaseem A, Snow V, Shekelle P, et al. Screening for HIV in health care settings: a guidance statement from the American College of Physicians and HIV Medicine Association. Ann Intern Med 2009;150:125–31.
[41] Kanwal F, Schnitzler MS, Bacon BR, et al. Quality of care in patients with chronic hepatitis C virus infection. Ann Intern Med 2010;153:231–9.
[42] Groom H, Dieperink E, Nelson DB, et al. Outcomes of a hepatitis C screening program at a large urban VA medical center. J Clin Gastroenterol 2008;42:97–106.

J Thromb Thrombolysis (2012) 34:143–163
DOI 10.1007/s11239-012-0700-3

Highlights from the IV International Symposium of Thrombosis and Anticoagulation (ISTA), October 20–21, 2011, Salvador, Bahia, Brazil

Renato D. Lopes · Richard C. Becker · L. Kristin Newby · Eric D. Peterson · Elaine M. Hylek · Christopher B. Granger · Mark Crowther · Tracy Wang · Antonio C. Carvalho · Otavio Berwanger · Roberto R. Giraldez · Gilson Soares Feitosa · Jorge Pinto Ribeiro · Eduardo Darze · Renato A. K. Kalil · Marianna Andrande · Fabio Villas Boas · Jadelson Andrade · Ana Thereza Rocha · Robert A. Harrington · Antonio C. Lopes · David A. Garcia

Published online: 17 March 2012

Abstract To discuss and share knowledge about advances in the care of patients with thrombotic disorders, the Fourth International Symposium of Thrombosis and Anticoagulation was held in Salvador, Bahia, Brazil, from October 20–21, 2011. This scientific program was developed by clinicians for clinicians and was promoted by three major clinical research institutes: the Brazilian Clinical Research Institute, the Duke Clinical Research Institute of the Duke University School of Medicine, and Hospital do Coração Research Institute. Comprising 2 days of academic presentations and open discussion, the symposium had as its primary goal to educate, motivate, and inspire internists, cardiologists, hematologists, and other physicians by convening national and international visionaries,

R. D. Lopes (✉) · R. C. Becker · L. K. Newby · E. D. Peterson · C. B. Granger · T. Wang · R. A. Harrington
Duke Clinical Research Institute, Duke University Medical Center, Box 3850, 2400 Pratt Street, Room 0311 Terrace Level, Durham, NC 27705, USA
e-mail: renato.lopes@duke.edu

R. D. Lopes · A. C. Carvalho · A. C. Lopes
Federal University of São Paulo, Paulista School of Medicine, São Paulo, SP, Brazil

E. M. Hylek
Boston University, Boston, MA, USA

M. Crowther
McMaster University, Hamilton, ON, Canada

O. Berwanger
Research Institute HCOR, Heart Hospital, São Paulo, SP, Brazil

R. R. Giraldez
São Paulo University Medical School, São Paulo, SP, Brazil

G. S. Feitosa
Bahiana School of Medicine and Public Health, Salvador, BA, Brazil

J. P. Ribeiro
Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil

E. Darze
Cardio-Pulmonar Institute, Salvador, BA, Brazil

R. A. K. Kalil
Federal University of Health Science of Porto Alegre, Porto Alegre, RS, Brazil

M. Andrande
Sirio Libanes Hospital, São Paulo, SP, Brazil

F. V. Boas
Spanish Hospital, Research Center, Salvador, BA, Brazil

J. Andrade
Bahia Hospital, Salvador, BA, Brazil

A. T. Rocha
Federal University of Bahia, Salvador, BA, Brazil

D. A. Garcia
University of New Mexico Health Sciences Center, Albuquerque, NM, USA

thought-leaders, and dedicated clinician-scientists. This paper summarizes the symposium proceedings.

Keywords Thrombosis · Antithrombotic therapy · Guidelines · Clinical research

Introduction

Venous and arterial thrombosis cause more than 50% of deaths in the developed world (http://www.cdc.gov/nchs/fastats/deaths.htm). Anticoagulants and antiplatelet drugs are the main tools used by physicians to prevent the formation of pathologic clots. During the first decade of the 21st century, dozens of clinical trials have been undertaken to evaluate promising new antithrombotic agents that offer the possibility of simpler treatment with either better efficacy, reduced toxicity, or both.

The Fourth International Symposium of Thrombosis and Anticoagulation was held in Salvador, Bahia, Brazil, from October 20–21, 2011; this congress offered its attendees the rare opportunity to network with each other while hearing about cutting-edge clinical research and discussing its implications for clinical practice with internationally recognized experts. This scientific program was developed for practicing clinicians in multiple specialties, and the meeting was endorsed by three major clinical research institutes: the Brazilian Clinical Research Institute, the Duke Clinical Research Institute of the Duke University School of Medicine, and Hospital do Coração Research Institute. It was also supported by the Brazilian Societies of Internal Medicine, Cardiology, Intensive Care Medicine, and Vascular Surgery, by the Latin American Group of Thrombosis and Hemostasis, and by the Anticoagulation Forum from the United States. The chairmen of the meeting were Dr. Renato D. Lopes and Dr. Richard C. Becker, both from Duke University School of Medicine and the Duke Clinical Research Institute, and Dr. David Garcia from the University of New Mexico.

After reading this summary of the symposium proceedings, we are confident the reader will agree that the symposium met its main goal: to educate, motivate, and inspire internists, cardiologists, hematologists, and other physicians to thoughtfully apply the best available evidence to the care of their patients with (or at risk for) thrombotic disease.

Platelet biology

Platelets are small cellular fragments devoid of a nucleus, derived from the megakaryocytes, with diameters ranging from 1.5 to 3.0 μm. In steady state, platelets assume a discoid shape that completely changes during activation. The half-life of platelets in circulation is approximately 8–12 days. Platelets contain large deposits of adenosine diphosphate (ADP) and adenosine triphosphate (ATP); thus, they have a high capacity for energy metabolism that is similar to that of smooth muscle cells. However, because platelets are devoid of a nucleus, they have limited ability for protein synthesis.

The main function of platelets is to ensure primary hemostasis. However, in several clinical scenarios (such as acute and chronic coronary syndromes and cerebrovascular diseases), platelets play a negative role and are considered the main elements responsible for the physiopathology of these serious diseases.

The platelet membrane is composed of proteins, carbohydrates, and lipids. Lipids represent 35% of the membrane composition and are anti-symmetrically arranged when the platelets are in steady state (not activated), with the negatively charged phospholipids arranged in the internal portion of the membrane. The external surface of the membrane is rich in receptors, among which are the glycoprotein complex (GP) Ib/V/IX that preferably binds von Willebrand factor (vWF); GP VI that strongly binds collagen; and GP IIb/IIIa that binds fibrinogen, allowing platelet aggregation. Also in the platelet membrane, proteins such as the p-selectin are expressed. These work as chemo-attractants for leukocytes.

Cytoplasmic organelles are also very important for platelet function. These include dense peroxisomes (responsible for lipid metabolism), mitochondria (oxidative metabolism), lysosomes (only released in response to very powerful stimulations; able to cause local injury), and dense granules and alpha granules. Dense granules contain calcium at high concentrations, as well as ADP, ATP, and serotonin. Alpha granules contain great variety of substances with pro-coagulant, mitogenic, and inflammatory functions (e.g., vWF, fibrinogen, cytokines, and platelet factor 4). The release of these granules varies according to platelet stimulation.

The platelet cytoskeleton is essentially undone and remodeled during platelet activation, changing from a discoid to a spherical shape. Also during platelet activation, phyllopods are formed. Contraction of the cytoskeleton contraction is one of the basic steps of platelet activation that allows secretion of dense and alpha granules.

Platelet function can be categorized into three actions: adhesion, activation, and aggregation. An initial trigger for platelet adhesion is vascular injury, with exposure of the subendothelial content (vWF and collagen). This vascular injury can be spontaneous or iatrogenic (e.g., during percutaneous coronary intervention [PCI]). Once vWF and collagen are exposed in the circulation, platelets initiate adhesion, through binding of the GP Ib/V/IX receptor to vWF and GP VI to collagen.

The degree of platelet activation depends on several vascular injury characteristics, such as: depth of the injury, vessel site, hematocrit level, flow speed at the site concerned, and vessel diameter. The activation process is initiated right after adhesion, and both continue to occur simultaneously. Four steps are fundamental during activation: mobilization of intracellular calcium (which works as the most powerful second messenger within the platelet); cytoskeleton contraction (with change in platelet shape); secretion of alpha and dense granules (with release of platelet agonists with autocrine and paracrine action; thus, enhancing the activation signal); and exposure of negatively charged phospholipids in the external portion of the membrane (with consequent activation of coagulation cascade and thrombin generation).

The more important platelet agonists are: thromboxane A2 (TP receptor), which stimulates initial activation and local vasoconstriction; ADP (acts on $P2Y_1$ and $P2Y_{12}$ receptors), which stimulates more stable activation; thrombin (acts on PAR 1 and 4 receptors), which primarily stimulates activation in pathological conditions and is considered the most powerful agonist; and collagen (GP Ib and GP VI), which stimulates activation.

Platelet aggregation is considered the final step in platelet response to injury and involves the conformational change of the GP IIb/IIIa receptor, which moves from a low-affinity steady state to a high-affinity activated state. The activated GP IIb/IIIa receptor binds fibrinogen, forming platelet-fibrinogen-platelet aggregates and a stable platelet plug.

Platelets interact directly with the coagulation system in many ways. The interaction can be physical, such as through exposure of negatively charged phospholipids on the external surface of the membrane, or chemical, through release of pro-coagulant granules. Platelets also interact similarly with inflammatory cells through exposure of p-selectin, leading to recruitment of leukocytes and exposure of the CD40 receptor from the surface of macrophages.

In short, platelets are important structures responsible for primary hemostasis, but play a negative role in several clinical scenarios. Platelet biology is very complex because platelets have to interact with other systems, including the coagulation cascade**.**

Measures of platelet function: are we ready to use them?

In the treatment of coronary disease, inhibition of platelet activation and aggregation is critical to the prevention of cardiovascular atherothrombotic outcomes. Clopidogrel, a $P2Y_{12}$ receptor antagonist, improves outcomes among patients with acute coronary syndrome (ACS) and after PCI. However, clopidogrel is a biologically inactive prodrug that requires several steps of metabolism for active effect. In part because of these activation steps, substantial inter-individual variability in pharmacodynamic response has been previously observed. Individual variation in response to antiplatelet therapies is in part predicated on intrinsic factors (such as genetic polymorphisms affecting absorption and/or metabolism of drug) and in part related to clinical factors such as patient non-compliance with therapies or drug–drug interactions.

There are several methods for assessment of platelet response to therapy. The current gold standard is light transmission platelet aggregometry, which involves introduction of a platelet agonist such as ADP with a light-based assay that ultimately assesses platelet aggregation. The requirement for high technical expertise, as well as substantial processing times to generate platelet-rich plasma, render this assay an unwieldy tool for clinical use. Point-of-care aggregation tests involving whole blood samples—such as the VerifyNow assay—employ the same principles in a cartridge-based fashion and have been validated against the gold standard. The vasodilator activated phosphorylation (VASP) assay measures intra-platelet phosphorylation in response to $P2Y_{12}$ receptor activation; higher VASP phosphorylation levels are observed with superior inhibition of the $P2Y_{12}$ receptor by agents such as clopidogrel. Yet, like light transmission aggregometry, this is a tool largely used for research purposes due to its technically demanding laboratory processes. Another platelet function testing modality, the multi-plate assay, examines the degree of platelet adhesion and aggregation on the sensors' surface by quantifying electrical resistance between the two central wires.

All of the above platelet function testing modalities have been shown to correlate with post-PCI outcomes. For VerifyNow, platelet reactivity unit levels > 235 are associated with increased risk of cardiovascular death, non-fatal myocardial infarction (MI), and stent thrombosis. The Do Platelet Function Assays Predict Clinical Outcomes in Clopidogrel-Pretreated Patients Undergoing Elective PCI study compared VerifyNow to traditional light transmission aggregometry and noted good correlation between these two tests in their ability to discriminate future thrombovascular outcomes. For VASP phosphorylation, low post-treatment response defined as platelet reactivity $> 50\%$ was associated with future risk of cardiovascular events. Similarly for the multi-plate assay, low responders were associated with a higher incidence of stent thrombosis.

The role of platelet function testing in tailoring therapy for individual patients remains to be elucidated. The multicenter Gauging Responsiveness With A VerifyNow

Assay—Impact on Thrombosis and Safety (GRAVITAS) trial focused on patients who had residual high-on treatment platelet reactivity measured by VerifyNow testing and randomized these patients to standard-dose clopidogrel (75 mg daily) versus higher-dose (150 mg daily). Unfortunately, no differences in clinical outcomes were noted between groups, which may in part be attributed to persistently high platelet reactivity even among the higher-dose group. While this study examined the impact of dose doubling, we look to studies such as the Double Randomization of a Monitoring Adjusted Antiplatelet Treatment Versus a Common Antiplatelet Treatment for DES Implantation, and a Interruption Versus Continuation of Double Antiplatelet Therapy (ARCTIC) trial, which will examine the role of tailored therapy using alternative agents, such as prasugrel or ticagrelor, that have been shown to provide more potent and consistent inhibition of platelet function.

One important question that has been raised is: what is the appropriate timing of platelet function testing? In a study by Campo et al., PCI patients had serial measurements at baseline and at one month. Among patients who were full responders at baseline, 4% became poor responders by 30 days; in contrast, among patients who were poor responders at baseline, 70% became full responders by one month. Further, patients who were poor responders both at baseline and at one month and those who were full responders at baseline but became poor responders at one month had worse ischemic outcomes by one year. These results suggest that the 30-day time point may be a more relevant time to test.

Another question is whether platelet function testing can be used to define a therapeutic window for antiplatelet therapies similar to the international normalized ratio (INR) for warfarin therapy. To date, the ability of existing platelet function tests to predict bleeding outcomes is limited. One modestly positive study by Sibbing et al. showed that bleeding was associated with an area under the curve $\leq$188 using multi-plate technology.

In summary, several testing modalities are currently available to assess on-treatment platelet response to antiplatelet therapies. While these tests provide important prognostic information for ischemic events, their potential for bleeding prediction appears limited. Platelet function testing may someday be helpful for therapeutic selection; however, further evidence is necessary.

Vitamin K antagonists: is this the beginning of the end?

For centuries, thrombosis has been recognized as a major pathological finding in many significant and often fatal clinical conditions. Parenteral anticoagulants, specifically unpurified heparin, led the way in the pharmacological treatment of thromboembolic diseases. Its main drawback was that it was not available for oral use. Shortly after, there came the anti-vitamin K oral anticoagulants—dicumarol and warfarin—which have been used widely since the 1950s for treatment and prevention of thromboembolic diseases, such as deep vein thrombosis (DVT) and pulmonary embolism (PE), and prevention of cerebral vascular embolism in conditions such as atrial fibrillation (AF), artificial cardiac valves, and ventricular thrombi.

Use of vitamin K antagonists was, from the very beginning, recognized as troublesome due to the very narrow therapeutic range of anticoagulation, which led to lack of protection when suboptimal and to hemorrhagic events when in excess of prescribed limits. Furthermore, many natural nutrients and medicines can interfere with the pharmacological action of vitamin K antagonists, requiring frequent monitoring of the anticoagulant status via repeated measurements of the prothrombin time. The INR (a standardized reporting method) must be maintained within a narrow therapeutic range to maximize the benefit of warfarin. Several observational studies and well-controlled trials have documented the challenges of long-term warfarin use. Among observational studies, the percent time in therapeutic range has ranged only from 50 to 55% and, in the controlled trials, from 58 to 65%.

All of these difficulties probably explain the under-use of oral vitamin K anticoagulants worldwide. The recent development of new oral anticoagulants with better pharmacological profiles and easier use has raised hopes that, in the near future, vitamin K antagonists will be replaced, assuming that the newer agents prove as successful in long-term surveillance as they have been demonstrated to be in relatively short-term trials.

New anticoagulants and new hematologic dilemmas

The development of novel anticoagulant medications has become a high priority for pharmaceutical companies. Enthusiasm for novel agents has resulted from the observation that about 50% of the population of the western world dies from either heart attack or stroke and that millions of people have AF, many of whom are inadequately treated using current oral anticoagulants. Additionally, our current armamentarium of anticoagulants (consisting predominantly of heparin, low-molecular-weight heparin [LMWH], oral vitamin K antagonists, and a selection of relatively infrequently used newer agents such as fondaparinux, hirudin, argatroban, and bivalirudin) are often perceived to be "old" and may have significant limitations that restrict their use. With the exception of the oral vitamin K antagonists, all of these medications are parenteral,

and many are expensive. Oral vitamin K antagonists are highly effective but have a slow onset and offset of action, large between-person variability in their dose requirements, are subject to food and drug interactions, and are complex to reverse. Despite these drawbacks, the oral vitamin K antagonists have been proven to reduce thromboembolism in a wide variety of settings, including patients with AF, mechanical heart valves, MI, and after orthopedic surgery. They are also effective for the secondary prevention of DVT, PE, and MI.

The "ideal" anticoagulant would be orally administered, have a rapid onset and offset of action, predictable pharmacokinetics and pharmacodynamics, a low propensity for food and drug interactions, be administered in fixed doses, be "reversible" in cases of bleeding, have a wide therapeutic window, and not require routine monitoring, but have a form of monitoring available should it be required. Oral vitamin K antagonists do not possess many of these characteristics.

Recent developments in the field of anticoagulants leave us at a crossroads; many clinicians are considering whether it is time to abandon the oral vitamin K antagonists. The rationale for reducing or eliminating use of oral vitamin K antagonists include eliminating the need for monitoring, having less variability in the dosing of the oral anticoagulant, and potentially reducing bleeding. However, eliminating oral vitamin K antagonists will be difficult. In some settings, they are the only proven therapy (e.g., patients with antiphospholipid antibody syndrome or those with mechanical heart valves). Additionally, oral vitamin K antagonists are inexpensive, they have 100% brand recognition internationally (thereby reducing the likelihood of medication errors), and the ability to monitor these drugs improves compliance. Finally, oral vitamin K antagonists can be rapidly reversed.

The development of novel anticoagulant medications has been facilitated by a comprehensive analysis of the coagulation cascade. Coagulation is initiated at sites of vascular injury when tissue factor binds with circulating activated factor VIIa. This complex converts factor X to factor Xa and factor IX to factor IXa. Factor Xa then acts in concert with factor Va to convert prothrombin to thrombin. Thrombin is the "engine" of coagulation. It converts fibrinogen to fibrin, facilitates a positive feedback loop leading to activation of coagulation, activates factor XIII, which cross-links fibrin stabilizing the clot, and has a number of other important roles in coagulation.

A clear understanding of the coagulation cascade has allowed the development of highly specific inhibitors of coagulation. Initially, these agents were developed using recombinant DNA technology modeled after naturally occurring anticoagulants. Perhaps the best example of this is the development of the hirudins, modeled after the anticoagulant present in the saliva of the medicinal leech. More recently, knowledge of the structure of the coagulation enzymes has allowed the development of molecules using computer-assisted design. These molecules are of low molecular weight, can be made orally bioavailable, and when carefully designed are highly specific to their enzyme target. Development of these agents has revolutionized the approach to anticoagulation.

Polypeptide drugs tested as inhibitors of coagulation include tissue factor pathway inhibitor, nematode anticoagulant peptide C2 (rNAPC2), active site-blocked factor VIIa, activated protein C, soluble thrombomodulin, and the hirudins. Low-molecular-weight inhibitors include rivaroxaban, apixaban, edoxaban, and dabigatran. Additional agents are under development.

In general, polypeptide drugs are falling out of favor, except for their use in acute situations such as ACS or in unstable patients with extensive thromboembolic disease. These drugs are used less and less frequently because of their high cost and need for parenteral administration. Furthermore, the effectiveness of many of these drugs has recently been questioned. For example, recombinant activated protein C was recently withdrawn from the worldwide market (http://www.hc-sc.gc.ca/ahc-asc/media/advisories-avis/_2011/2011_142-eng.php, accessed November 14, 2011). Some of these drugs (e.g., rNAPC2) are being studied for other indications, such as the treatment or prevention of cancer.

Recent research has focused on inhibition of factor Xa or thrombin, given their seminal roles in coagulation. Rivaroxaban, apixaban, and edoxaban are all highly active, orally bioavailable inhibitors of coagulation. Each has been extensively tested in diverse clinical situations, and each holds promise as a therapy for the prevention and treatment of both venous and arterial thrombosis. Dabigatran is the only thrombin inhibitor currently available; although it has reduced bioavailability, it has demonstrated its effectiveness as an agent for the prevention and treatment of thrombosis.

Novel agents should not be regarded as a panacea; although they address many of the perceived concerns with novel anticoagulants, they have their own limitations. These include cost, a lack of familiarity among the medical community (leading to the likelihood of medical error), a much less broad set of indications (given their development in a more restrictive regulatory environment), and a lack of antidotes or reversibility. Although there has been early work on the development of specific antidotes for both the direct Xa inhibitors and dabigatran, none of these agents has reached even early-phase clinical trials in patients with active bleeding.

Given the excitement surrounding novel anticoagulant medications, it is very likely that a great deal more research

describing their utility and toxicities will be undertaken in the coming years.

Antiplatelet agents under development

The evolution of antiplatelet therapy over the past decade has witnessed clear advances, with the development of increasingly potent and response-consistent $P2Y_{12}$ receptor antagonists, including prasugrel, and more recently ticagrelor, a non-thienopyridine agent that may also offer benefit through prolonged inhibition of adenosine re-uptake by erythrocytes. A wealth of information from phase three clinical trials highlights the broad potential of new-generation antiplatelet agents but also underscores the importance of patient selection, aspirin dosing, and uncommon yet potentially life-threatening hemorrhagic complications involving the gastrointestinal tract and brain. These adverse events serve as a reminder that platelets play an important role in hemostasis and the maintenance of vascular integrity, including the blood–brain barrier.

Future investigations will likely focus on strategies and technologies to optimize patient-centered therapies and platelet antagonists that attenuate thrombosis while preserving hemostatic potential and vascular reparative capacity.

Anticoagulation in ACS patients managed invasively: a time for change or a time of choice?

Earlier studies demonstrated that the use of enoxaparin in non–ST-segment elevation ACSs (NSTE ACS) patients managed conservatively reduced the rates of death or MI by approximately 20% both at 8 and 42 days compared with unfractionated heparin (UFH). More recent studies, however, indicated that an invasive strategy with early catheterization and angioplasty was associated with improved outcomes compared with a conservative approach. When enoxaparin was compared with UFH in the setting of an early invasive strategy, similar efficacy outcomes, regardless of anticoagulation therapy, were seen. In contrast, patients treated with enoxaparin had a significant 30% increase in severe bleeding as evaluated by the TIMI scale.

Recently, a great deal of evidence has suggested that bleeding is a major determinant of clinical outcomes in ACS patients. Both moderate and important bleeding are associated with worse outcomes after adjustment for potential confounders. In a pooled analysis of 26,452 patients with ACS, severe bleeding increased more than five times the odds of 30-day mortality or MI.

Drugs recently developed, such as factor Xa inhibitors and direct thrombin inhibitors, have an improved safety profile. In the Fifth Organization to Assess Strategies in Acute Ischemic Syndromes (OASIS-5) trial, the indirect Xa inhibitor fondaparinux reduced major and minor bleeding by more than 50% in NSTE ACS patients. This resulted in a significant 11% reduction in mortality at six months. Similarly, the thrombin blocker bivalirudin reduced severe bleeding by almost 40% in ST-segment elevation MI patients, leading to a 30% reduction in all-cause mortality. This benefit extended to three years of follow-up.

The results of these trials indicate that safer anticoagulant drugs with similar efficacy profiles are currently available and that new treatment combinations may reduce mortality. Therefore, it is a time for change to the newer agents in ACS. At this point, it is very important that the guidelines and future studies focus on these new therapeutic options; clinicians can benefit from more evidence when choosing the best clinical setting for each particular agent.

Oral anticoagulants after ACS

Current guidelines recommend dual antiplatelet therapy after ACS, but the risk of recurrent ischemic events remains elevated in these patients. Meta-analyses of clinical trials on the addition of warfarin to aspirin after ACS indicate that this intervention is associated with a reduced rate of ischemic events at the cost of increased risk of major bleeding. Moreover, warfarin therapy requires frequent monitoring and is also associated with food and drug interactions. With the development of oral agents that directly inhibit thrombin or drugs that are direct inhibitors of factor Xa, a new opportunity for secondary prevention after ACS has emerged. These agents have predictable dose-dependent pharmacokinetics and pharmacodynamics and, therefore, do not require frequent monitoring. Some have also been shown to be effective and safe in the management of AF. Dabigatran, a direct inhibitor of thrombin, was compared with placebo on top of dual antiplatelet therapy in the Reduction by Dutasteride of Clinical Progression Events in Expectant Management (REDEEM) phase two trial in 1861 patients. The administration of newly developed direct factor Xa inhibitors on a background of single or dual antiplatelet therapy has also been evaluated by phase two trials. In the Rivaroxaban in Combination With Aspirin Alone or With Aspirin and a Thienopyridine in Patients With ACSs (ATLAS ACS TIMI-46) trial, different doses of rivaroxaban were compared with placebo in 3491 patients. Darexaban was also evaluated in the dose-ranging Study Evaluating Safety, Tolerability and Efficacy of YM150 in Subjects With ACSs (RUBY) trial with 1279 patients, and apixaban was evaluated in the Apixaban for Prevention of Acute Ischemic

Events (APPRAISE) dose-ranging trial with 1714 patients. The results of these phase two trials in patients after ACS are remarkably consistent, by showing a dose-dependent increase in major bleeding with little or no significant effect on the reduction of cardiovascular events. In agreement with these findings, a phase three trial, APPRAISE-2, with 7392 patients, was prematurely terminated because of an increase in major bleeding events with apixaban in the absence of a counterbalancing reduction in recurrent ischemic events.

The role of factor Xa inhibitors in the secondary prevention after ACS became even more complex after the results of the ATLAS-ACS 2 TIMI 51 trial, which evaluated more than 15,526 patients randomized to placebo or to two doses of rivaroxaban. Results of this highly anticipated trial demonstrated that ACS patients receiving standard therapy, including dual antiplatelet therapy, may benefit from the addition of the factor Xa inhibitor rivaroxaban, although at the cost of some additional bleeding complications. Both rivaroxaban doses reduced the primary end point of cardiovascular death/MI/stroke at the cost of increased bleeding rates. The 2.5-mg twice-daily dose had the better benefit/risk balance, due to lower bleeding risk, than the 5-mg twice-daily dose. Surprisingly, the lower dose of rivaroxaban resulted in a significant reduction in death from cardiovascular causes (2.7% vs. 4.1%, $P = 0.002$) and in all-cause mortality (2.9% vs. 4.5%, $P = 0.002$). These benefits were not observed in higher-dose rivaroxaban, and the difference between the two doses of rivaroxaban was significant. Rivaroxaban-treated patients experienced more major bleeding and intracranial hemorrhage than controls, but without a significant increase in fatal bleeding. It is possible that the addition of very-low-dose anticoagulation with rivaroxaban may represent a new treatment strategy in patients with a recent ACS. However, it is important to note that the ATLAS-ACS 2 trial had relatively small percentages of elderly patients, female patients, and patients with impaired renal function, suggesting that the results may not be entirely replicated with higher-risk patients in the real world.

In summary, as mentioned in the editorial by Matthew Roe and E. Magnus Ohman, "a new era of secondary prevention after an ACS has begun and will be characterized by the need to balance ischemic versus bleeding risks when selecting the type, number, and duration of antithrombotic therapies for individual patients."

New antiplatelet agents in ACS patients: how should we choose?

Treatment options for patients with ACS continue to expand. Most recently, two new potent oral inhibitors of the platelet $P2Y_{12}$ receptor, prasugrel and ticagrelor, were found in randomized clinical trials to be superior to clopidogrel in preventing death, MI, or stroke in patients presenting with ST-segment elevation MI and NSTE ACS and are now available for clinical use. However, their availability only adds to the complexity of treatment selection. Considering combinations of oral and intravenous (IV) antiplatelet agents, anticoagulants, and their use and timing relative to the use and timing of PCI, there are at least 144 different possible combinations for treatment of an individual patient. Given this complexity, a number of factors should be considered in selecting therapy.

First and foremost, the evidence supporting efficacy and safety of the agent must be considered. In the Clopidogrel in Unstable Angina to Prevent Recurrent Events (CURE) trial, clopidogrel reduced the risk of death, MI, or stroke relative to aspirin alone by 20% with acceptable incremental bleeding. However, the CURE trial was conducted prior to the current era of invasive treatment for NSTE ACS and did not include ST-segment elevation MI patients. In addition, clopidogrel levels and platelet responsiveness to clopidogrel are highly variable across patients, in part related to polymorphisms in CPY2C19, an enzyme in the cytochrome P450 system responsible for converting clopidogrel from a pro-drug to its active form. Unfortunately, there is no evidence to date that increasing the dose of clopidogrel is effective in reducing clinical events in patients with reduced function polymorphisms of CYP2C19 or in reducing death, MI, or stroke in ST-segment elevation MI and NSTE ACS patients overall.

Like clopidogrel, prasugrel is a thienopyridine and administered as a pro-drug, but is more potent than clopidogrel in inhibiting platelet function with less variability in response across individuals. Additionally, prasugrel is less susceptible to the effects of CYP2C19 polymorphisms than clopidogrel.

Ticagrelor is a non-thienopyridine, reversible inhibitor of the $P1Y_{12}$ receptor that is more potent than clopidogrel and is administered as the active drug. Briefly, compared with clopidogrel in patients with both NSTE ACS and ST-segment elevation MI, prasugrel (administered after coronary anatomy was known and PCI was planned) and ticagrelor (administered as upstream therapy) reduced the risk of death, MI, or stroke by 19 and 16%, respectively, over treatment of approximately 12 months. Importantly, treatment with ticagrelor resulted in a significant 21% reduction in cardiovascular mortality. Efficacy results were consistent across major subgroups, with a possible enhanced benefit of prasugrel among diabetic patients and reductions in stent thrombosis with both agents. There was a significant treatment-by-region interaction in the Platelet Inhibition and Patient Outcomes (PLATO) trial, such that the point estimate for treatment effect favored clopidogrel

in North America. Subsequent analyses suggested that aspirin dose explained the majority of the difference in treatment response by region, and ticagrelor now carries a Food and Drug Administration (FDA) black box warning for use only with low-dose aspirin (<100 mg).

Although there was a higher rate of major bleeding with ticagrelor compared with clopidogrel, there was no increase in intracranial or life-threatening bleeding. However, prasugrel resulted in a 32% higher rate of major bleeding than clopidogrel, which included significant increases in life-threatening and fatal bleeding and an increase in intracranial hemorrhage in patients with prior transient ischemic attack or stroke. Net clinical benefit was also not favorable for prasugrel in patients over age 75 years and those with body weight <60 kg. Therefore, prasugrel has an FDA black box warning restricting use in these subgroups.

A practical consideration in selecting treatment is cost, particularly with the expectation that clopidogrel will become available in generic form in May 2012. Present pricing for consumers in the United States is approximately $200 per month for clopidogrel, $215 per month for prasugrel, and $260 per month for ticagrelor. Compliance is also a practical concern. Both clopidogrel and prasugrel were developed for once-daily dosing, but ticagrelor was developed to be given twice daily, which may affect compliance in some patients. Additionally, ticagrelor causes symptomatic dyspnea in some patients, which may lead to lower compliance.

Selection among the new antiplatelet agents and clopidogrel should consider local practice patterns, including such things as what agents are available in institutional formularies and in local pharmacies and how familiar physicians at the treating facility are with the properties of the agents and their potential benefits and risks. Finally, a systematic approach is recommended for oral antiplatelet treatment selection that considers these factors, as well as local interventional and referral patterns, and that simplifies the approach to selection of all antithrombotic therapy for treatment of ACS at an institution.

Bleeding and mortality in patients with ACS

Bleeding occurs commonly during the treatment of ACS, an observation that has been made in both clinical trials as well as in observational registries of community practice. However, the incidence of bleeding depends on the definition applied, and unfortunately varying definitions have been applied historically across trials. For example, in the TIMI trials, major bleeding is defined as fatal or life-threatening bleeding, intracranial hemorrhage, hemoglobin drop ≥5 g/dL, or hematocrit drop ≥15%. In contrast, the Acute Catheterization and Urgent Intervention Triage Strategy definition of major bleeding includes intracranial, retroperitoneal, intraocular, or access-site hemorrhage requiring surgical intervention, hematomas ≥5 cm in diameter, as well as more conservative hemoglobin drops depending on whether an overt source of bleeding is observed. The use of blood product transfusions has been variably incorporated into these definitions as well and contributes further to the variation and incidence of bleeding due to varying institutional thresholds for transfusion. The Bleeding Academic Research Consortium (BARC) assembled a working group to provide standardized bleeding definitions for cardiovascular clinical studies. This group acknowledged several challenges in creating a universal bleeding definition but proposed a consensus classification for bleeding evaluation or treatment. Type II bleeding includes clinically overt signs of bleeding that are actionable but do not meet criteria for other BARC bleeding. Type III is clinical, laboratory, and/or imaging evidence of bleeding with specific health care provider response. Type IV includes coronary artery bypass graft (CABG)-related bleeding, and Type V is defined as fatal bleeding.

Bleeding has important prognostic complications as it has been associated with an increased longitudinal risk of mortality. The severity of bleeding correlates with worse outcomes. Bleeding is costly, not only because it prolongs length of hospitalization, but also because it is associated with additional diagnostic and treatment interventions. Bleeding can also result in decreased use of evidence-based therapies. In a study of patients with ACS, patients who developed bleeding complications were less likely to be discharged on antiplatelet therapy such as aspirin and thienopyridine. Six months after a bleeding event, these patients remained less likely to receive these evidence-based therapies. Even nuisance bleeding is associated with a high rate of antithrombotic treatment discontinuation.

Bleeding also often leads to the transfusion of blood products. Tremendous variation in transfusion practices exists across health care organizations. However, a common theme is that older patients, female patients, and patients with renal insufficiency are more likely to receive transfusion therapy. Transfusion in the ACS setting has been independently associated with worse outcomes. Appropriate thresholds for transfusion have not been well established within the cardiovascular arena. The only randomized clinical trial on this topic was conducted by the Canadian Clinical Trials group in 1999, which randomized 838 critically ill patients to a liberal versus restrictive transfusion strategy. No difference in 30-day mortality was noted between transfusion strategies, although there was a suggestion of benefit for the more liberal strategy among patients with coronary disease.

There is increased focus on bleeding avoidance strategies. First and foremost would be the prediction of bleeding risk. The Can Rapid Risk Stratification of Unstable Angina Patients Suppress Adverse Outcomes with Early Implementation of the American College of Cardiology/American Heart Association Guidelines (CRUSADE) bleeding risk score was developed and validated in a cohort of patients with NSTE ACS, and it identified eight clinical factors that are associated with increased bleeding risk: female sex, prior history of diabetes, heart failure, peripheral vascular disease, or clinical and laboratory evidence of risk including lower creatinine clearance, higher heart rate on admission, lower systolic blood pressure on admission, and lower admission hematocrit. A key first step to bleeding avoidance is the appropriate dosing of antithrombotic therapies. Patients who are older or who have worsening renal function are particularly vulnerable to excess antithrombotic dosing. Appropriate dosing, particularly in these high-risk patients, may alleviate the relationship between these risk characteristics and increased bleeding risk. The use of safer antithrombotics may also be considered. Bivalirudin, for example, has been well studied in both the ST-elevation and NSTE ACS settings and has been shown to have efficacy comparable with the traditional combination of heparin and GP IIb/IIIa inhibitors but an enhanced safety profile with reduced bleeding risk. Arterial access for cardiac catheterization may also play a role in bleeding avoidance. The recent Radial Versus Femoral Access for Coronary Angiography and Intervention in Patients with Acute Coronary Syndromes (RIVAL) study compared clinical outcomes between patients who underwent radial versus femoral arterial access. Although the primary outcome of death, MI, stroke, and non-CABG-related major bleeding at 30 days was not significantly different between groups, secondary outcomes, including the occurrence of large hematomas or major vascular access site complications, were substantially reduced with radial access.

In summary, bleeding is an important event associated with increased mortality among patients with ACS, and bleeding avoidance should be prioritized while reducing ischemic risk with antiplatelet interventional therapies. Strategies that achieve this balance are associated with improved survival.

Thromboprophylaxis in critically ill patients

Critically ill patients have an increased risk of DVT due to their acute illness, procedures such as central venous catheterization, and immobility. Among patients in the intensive care unit (ICU), DVT is an important problem because thrombus propagation and embolization can lead to potentially fatal PE. The effects of thromboprophylaxis with LMWH compared with UFH on venous thromboembolism (VTE), bleeding, and other outcomes were uncertain in critically ill patients. To address this question, The Prophylaxis for Thromboembolism in Critical Care (PROTECT) trial (NCT00182143) was planned.

PROTECT was a randomized, stratified, concealed international trial comparing subcutaneous injection of UFH 5000 IU or the LMWH dalteparin 5000 IU once daily plus once-daily placebo for the duration of the ICU stay. The objectives of PROTECT were to examine, among medical-surgical critically ill patients, the effect of the LMWH versus heparin on the primary outcome of proximal leg DVT and the following secondary outcomes: DVT elsewhere, PE, any venous thromboembolism (DVT or PE), and the composite of VTE or death, bleeding, and heparin-induced thrombocytopenia. Patients were followed up to death or hospital discharge. Venous thromboembolism events were included after ICU discharge. All patients, families, clinicians, research personnel, outcome adjudicators, and the trial biostatistician were blinded to allocation. Data were analyzed according to the intention-to-treat principle.

The main results of the PROTECT trial suggested that there was no significant between-group difference in the rate of proximal leg DVT, which occurred in 96 of 1873 patients (5.1%) receiving dalteparin versus 109 of 1873 patients (5.8%) receiving UFH (hazard ratio in the dalteparin group, 0.92; 95% confidence interval [CI], 0.68–1.23; $P = 0.57$). The proportion of patients with pulmonary emboli (a key secondary outcome measure) was significantly lower with dalteparin (24 patients, 1.3%) than with UFH (43 patients, 2.3%) (hazard ratio, 0.51; 95% CI, 0.30–0.88; $P = 0.01$). There was no significant between-group difference in the rates of major bleeding (hazard ratio, 1.00; 95% CI, 0.75–1.34; $P = 0.98$) or death in the hospital (hazard ratio, 0.92; 95% CI, 0.80–1.05; $P = 0.21$).

Does aspirin have a role in venous thromboembolism prevention?

Patients undergoing orthopedic surgery are at high risk for VTE. Anticoagulant agents (e.g., LMWH or fondaparinux) significantly reduce the risk of VTE after orthopedic surgery, but the role of aspirin in VTE prevention has been controversial, in part because platelets are thought to be less important than fibrin in venous thrombosis. Guidelines published by the American Academy of Orthopedic Surgeons recommend aspirin as one of several acceptable VTE prevention strategies for the majority of patients undergoing major orthopedic procedures. In contrast, the eighth edition of the American College of Chest Physicians'

Evidence-based Guidelines on Antithrombotic Therapy strongly recommends that aspirin *not* be used for VTE prophylaxis.

A recent systematic review of all evidence relevant to aspirin suggests that aspirin is almost certainly more effective than placebo for the prevention of VTE following orthopedic surgery. The most important data come from two sources: a meta-analysis published by the Anti-Platelet Trialists' Coalition (APTC) in 1994 and the Pulmonary Embolism Prevention (PEP) study published in 1999. The APTC meta-analysis included data pooled from almost 9000 patients who had participated in VTE prevention trials following orthopedic surgery, as well as other clinical settings. Although the results indicated that, compared with placebo, aspirin reduced the rate of PE by more than 50%, the validity of the findings was questioned because many of the trials from which these data were abstracted had serious methodologic flaws (e.g., open-label design, sub-optimal DVT detection methods, non-uniform antiplatelet drugs and doses across studies).

The PEP trial was undertaken to address the fact that many physicians (especially non-surgeons) had rejected the APTC results. Over 17,000 patients in five countries were randomized to receive either aspirin 160 mg or placebo, started pre-operatively and continued for 35 days. The majority of patients underwent surgery to repair a hip fracture, but about one quarter of participants had an elective total hip arthroplasty. This well-designed, very large randomized trial with almost 100% follow-up for clinical end points found that aspirin reduced the risk of fatal and non-fatal PE; the relative effect was very similar to that seen for aspirin in the APTC meta-analysis. Based on data from PEP and APTC combined, it appears that, compared with placebo, aspirin could prevent 5–10 PEs per 1000 patients treated. Aspirin use following major orthopedic surgery would be expected to increase the number of patients who require transfusion by approximately 3–6 per 1000. Although a few direct comparisons between aspirin and anticoagulant agents in this setting have been published, the available evidence is too sparse and too inconsistent to draw definitive conclusions about the net clinical benefit (or harm) of anticoagulants versus aspirin in this setting. Pending further study, it seems reasonable to conclude that aspirin should have *some* role as a VTE prevention strategy after orthopedic surgery; without more data, however, we can expect ongoing disagreement about the patients for whom it would be most appropriate.

Should patients with cancer receive primary VTE prophylaxis?

The association between cancer and thromboembolic phenomena has been recognized since 1865. It was first described by Armand Trousseau (1801–1867) as a sign of occult pancreatic cancer, but there is wide variation in the relative risk of VTE in different cancers. The presence of an active cancer should be one of the leading risk factors recognized by physicians when assessing VTE risk during hospitalization. In the algorithm used for electronic alert to prevent VTE created by Kucher and colleagues, the presence of cancer carries a score of 3, while a score ≥ 4 reflects a high thrombotic risk. Some other scores evaluate the VTE risk in patients with cancer and on chemotherapy. The most important single message is that physicians should perform a systematic evaluation of each cancer patient, taking into account both factors linked to the patient (such as age, type and stage of the disease, previous history of VTE, known thrombophilia), as well as factors linked to the treatment (type of chemotherapy, hormonal therapy or an anti-angiogenic agent, such as thalidomide in association with corticosteroids, immobilization during hospitalization, and surgery).

VTE is one of the most frequent complications in cancer patients (4–20%). The risk of VTE is 3–5 times higher in cancer patients undergoing surgery than in those without cancer, and, as a consequence, among hospitalized cancer patients who die, one in every seven do so from PE. Also, when cancer patients develop a thrombotic event, they have a three-fold increased risk of recurrence, even years after the first VTE episode. Compared with similar patients without malignant disease, cancer patients have twice the incidence of bleeding during anticoagulant treatment. Furthermore, the development of VTE is independently associated with lower survival rates. At the same time, there is growing evidence that the use of anticoagulants lowers the risk of death; however, this hypothesis requires further research.

With regard to preventing thromboembolic disease in cancer patients, the patients who have been studied most often are those undergoing surgical intervention for their cancer. Low-dose UFH and LMWH are effective in preventing both DVT and fatal PE in general and in oncological surgical patients undergoing laparotomy. It has been demonstrated that cancer itself is a risk factor for the development of perioperative bleeding complications independent of pharmacological thromboprophylaxis type.

Cancer patients admitted to the hospital but not undergoing surgical intervention should be treated as other acutely ill hospitalized medical patients and provided with thromboprophylaxis when appropriate. The use of routine anticoagulation for patients with central venous catheters is no longer recommended. Although older studies suggested that the use of low-dose vitamin K antagonist or LMWH was associated with a benefit in reducing the frequency of thrombosis associated with central catheters, some more contemporary studies with the same agents fail to demonstrate a benefit.

At least three major guidelines for thromboprophylaxis in cancer patients have been published in recent years—the European Society for Medical Oncology, the American Society of Clinical Oncology, and the American College of Chest Physicians, 8th edition—and may help physicians in their clinical practices. Some of the key points about primary prophylaxis in cancer patients are as follows:

- LMWH, low-dose UFH, and vitamin K antagonist are not routinely indicated to prevent catheter-related thrombosis or during chemotherapy, if patients are ambulatory, except in multiple myeloma patients receiving thalidomide or lenalidomide or dexamethasone.
- Routine use of VTE prophylaxis with low-dose UFH or LMWH or fondaparinux in cancer patients undergoing medium and large surgical procedures is recommended.
- VTE prophylaxis should be maintained for at least 7–10 days and considered for up to 28 days in curative pelvic and abdominal cancer procedures.
- Routine use of prophylaxis should be considered in cancer patients hospitalized with an acute medical illness.

Managing anticoagulation in patients undergoing surgical procedures: diminishing bleeding and ischemic risks

Excessive bleeding leads to early instability and postoperative complications, and blood transfusion is clearly related to late mortality after cardiac surgery. Measures to prevent bleeding and avoid blood transfusion are very important to improve surgical outcomes. Preoperative anticoagulation and antiplatelet therapy comprise one of the six major risk factors for prediction of surgical bleeding.

Consequently, there are two situations in which coagulation status must be managed prior to surgical procedures to diminish bleeding and ischemic risk: in patients using an anticoagulation agent and in those receiving dual antiplatelet therapy. Aspirin is not a problem anymore. Although in the past it was recommended to cease aspirin use 3–5 days prior to surgery, this is no longer the case as guidelines now recommend that aspirin be administered prior to CABG because it has been shown to be related to better surgical outcomes. Thus, aspirin may be avoided only before special surgical procedures such as ophthalmic or intra-cerebral procedures.

Anticoagulants

Warfarin is by far the most widely used anticoagulant. In some regions, fenprocoumon is preferred because of its longer half-life. Newer agents, such as dabigatran, apixaban, and rivaroxaban, for example, have not yet been completely incorporated into the routine clinical armamentarium. Anticoagulants are mostly used for preventing arterial and pulmonary embolism in atrial fibrillation, for implanted mechanical valves, DVT, and PE. The main advantages of the newer anticoagulants are that they do not require monitoring and have a shorter half-life, which could facilitate management around the time of a surgical procedure. But there is very limited published evidence regarding those new drugs in surgery.

Warfarin might not be interrupted for low-risk endoscopic procedures. For patients at high risk for thrombosis who undergo procedures with low bleeding risk, it is recommended to stop use 3–5 days preoperatively. When the thrombosis risk is high and there is also high bleeding risk, it is recommended to stop warfarin 3–5 days preoperatively and monitor the INR. In the presence of an implanted mechanical heart valve, there is need to employ a bridge therapy prior to surgery, interrupting warfarin and introducing unfractionated or low-molecular-weight heparin until the moment of the procedure, and restarting warfarin the day after surgery. In this setting, a practical approach to perioperative anticoagulation would be: Discontinue warfarin 5 days preoperatively if INR is between 2 and 3 and for 6 days if INR is between 3 and 4.5. Observe that aging is associated with a slow resolution of anticoagulant effect.

Antiplatelet drugs

As already mentioned above, earlier guidelines recommended interruption of aspirin use some days prior to any operation. This is no longer true, especially in cardiac surgery, because with modern management, bleeding is seldom related to aspirin administration. Current guidelines do not recommend aspirin interruption, and the recently published American College of Cardiology/American Heart Association guideline for CABG surgery includes a Class I recommendation for the administration of aspirin in patients who are not taking it and who are undergoing CABG because it is related to better surgical outcomes.

Thienopyridines are related to postoperative bleeding and should be avoided or require specific management during the perioperative period. Because there are some different characteristics between the two most commonly used drugs of this class, they will be commented on separately below:

Clopidogrel

This is the most widely used ADP $P2Y_{12}$ inhibitor. Because its effect on platelets is irreversible, its use must be interrupted 5–7 days preoperatively, providing sufficient time for the platelet population to be renewed. The impact of exposure to clopidogrel in patients with ACS requiring coronary artery

bypass surgery was studied in a multicenter analysis for the end points of reoperation, major bleeding, and length of hospital stay. It was found that the adjusted risk for reoperation was 9.80 (95% CI: 2.18–43.95; $P = 0.01$) in the clopidogrel patient group. Interestingly, the management of patients undergoing CABG on clopidogrel seems to be improving, as bleeding and associated mortality have been reduced in recent reports as compared with the last two decades. There is variability in the patient's response to clopidogrel, due to genetic characteristics. For this reason, some would argue that there is a place for point-of-care testing to evaluate its action in the individual patient, although there is no formal recommendation for such testing. More studies to clarify the role of these tests in this setting are warranted.

Prasugrel

As with clopidogrel, this drug irreversibly binds the ADP $P2Y_{12}$ receptor. It has not been used as frequently as clopidogrel because it was more recently approved for clinical use. The Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel–Thrombolysis in Myocardial Infarction (TRITON–TIMI) found it to be rather efficacious despite a higher bleeding tendency in ACS. In patients receiving CABG, however, prasugrel was related to a 21.9% incidence of surgical bleeding, compared with 4.1% related to clopidogrel. It is reasonable to recommend stopping prasugrel for at least 7 days prior to CABG and to take special precautions to prevent perioperative bleeding.

Ticagrelor

This drug has some advantages over the above-mentioned thienopyridines because its effect is reversible and diminishes within 48 h, so that major surgical procedures can be done without fear of excessive bleeding. In the PLATO trial for ACS, ticagrelor was related to reduction of vascular death and MI but not stroke, compared with clopidogrel. There is little clinical evidence regarding the impact of ticagrelor on surgical outcomes. In a recently published analysis comparing ticagrelor with clopidogrel in the PLATO trial, among patients receiving CABG for whom ticagrelor/placebo was to be withheld for 24–72 h and clopidogrel/placebo for 5 days preoperatively, there was reduced cardiovascular and total death without an increase in major bleeding in the ticagrelor group.

Perioperative management in patients under dual antiplatelet therapy

As aspirin can be safely maintained during the surgical period, management is focused on the second antiplatelet drug. Taking clopidogrel as the paradigm drug and adjusting management for the others drugs, general measures can be summarized as below.

It is recommended as Class I in Society of Thoracic Surgeons (STS)/Society of Cardiovascular Anesthesiologists (SCA) guidelines to stop medications that inhibit the platelet $P2Y_{12}$ receptor before operation to decrease bleeding events. The timing of discontinuation depends on the pharmacodynamic half-life for each agent, as well as the potential lack of reversibility.

European Society of Cardiology/European Association for Cardio-Thoracic Surgery guidelines state that, for preoperative management of patients treated with dual antiplatelet therapy and considered for cardiac and non-cardiac surgery, one should proceed as follows: (1) emergent case: proceed to surgery; (2) semi-elective and urgent: "case-by-case" clinical decision; (3) elective: wait until completion of mandatory dual antiplatelet regimen. The decision-making process should balance the risk of thrombosis against the risk of bleeding, leading to continued use of aspirin + clopidogrel in low bleeding risk cases and continued use of aspirin but stopping of clopidogrel or even of both drugs in high bleeding risk cases.

The above-mentioned large meta-analysis on safety of clopidogrel being continued until CABG in ACS proposes an algorithm: (1) For stable elective CABG with no drug-eluting stent: stop clopidogrel for >5 days; (2) For stable elective CABG in the presence of drug-eluting stent implanted for <1 year: consider operating on clopidogrel or switch to IV tirofiban plus heparin as a bridge to surgery; (3) For non-elective CABG in emergent or urgent ACS: operate on clopidogrel, unless it is a reoperation, there is a bleeding disorder, or troponin is negative, favoring a multidisciplinary decision.

Lysine analogues, epsilon-aminocaproic acid, and tranexamic acid are very useful as they reduce total blood loss and decrease the need for blood transfusion during cardiac procedures and are indicated for blood conservation as Class I (Level A) in STS/SCA guidelines. Their effect is beneficial, even in routine CABG cases, for reducing postoperative thorax blood drainage. It might be advisable and is accepted in some centers for emergency ACS patients facing an intervention to avoid antiplatelet drugs until an anatomical coronary arteries diagnosis is obtained for deciding to proceed with a percutaneous intervention or surgery, or even excluding a non-coronary cause, such as aortic dissection, maintaining free use of aspirin and heparin.

Finally, some practical management tips based on surgical experience and previous literature are listed below:

- If possible, delay surgery for 3–5 days (clopidogrel or prasugrel) or 2–3 days (ticagrelor) if the patient is

relatively stable, balanced against severity and instability.
- If the patient is stable but the lesion is critical, give IV UFH until the effect of the antiplatelet drug wears off.
- No prophylactic preoperative transfusion of blood products.
- Do not stop heparin and give a full dose of heparin before bypass.
- Use tranexamic acid 10 mg/kg before surgical incision, plus 10 mg/kg on beginning of bypass.
- Generously use platelet transfusion after administration of protamine if diffuse bleeding, 0.2 units/kg.
- In the post-operative period, proceed to judicious surgical re-exploration, platelets infusion, and tranexamic acid when needed.
- Use reduced hemo-dilution under normothermia, modified ultrafiltration; minimize cardiopulmonary bypass circuits and priming.
- Employ good operative technique, including topical hemostasis, plus sealants or biological glues.
- Improve blood salvage methods and neutralize heparin with protamine on a titration basis.

Venous thromboembolism in heart failure patients: how should we manage this special population?

Venous thromboembolism and congestive heart failure (CHF) are among the most commonly encountered medical conditions, particularly in the elderly and hospitalized population. It is well recognized that CHF patients develop a hypercoagulable state that involves abnormalities in all three components of the Virchow's triad, which places these patients at an increased risk for VTE. Unfortunately, VTE has generally been regarded to be an end point of secondary importance in large, randomized CHF trials, and therefore, its actual incidence in this population has not been well studied.

Multiple autopsy studies have confirmed a very high prevalence of VTE in CHF patients. A case-controlled study has also shown that CHF is an important risk factor for VTE in ambulatory outpatients and that the risk progressively increases with worsening of the ejection fraction. Large, randomized CHF trials have suggested in retrospective analyses that the incidence of PE in mildly symptomatic ambulatory patients is probably low (<1%/year). On the other hand, hospitalized CHF patients with impaired mobility have an incidence of VTE as high as 21%, although most of the events are asymptomatic distal DVT. In a recent prospective study involving severe CHF patients admitted to a coronary care unit, PE was diagnosed in 9.1% during the hospitalization period, despite adequate prophylaxis in 70% of the cases. It is also important that PE patients with CHF have a higher mortality than those without CHF, and that PE is an independent predictor of death and rehospitalization in CHF patients.

The diagnosis of VTE in CHF patients is an increasingly frequent and challenging problem. There is a substantial overlap in symptoms and signs of both conditions, and some of the diagnostic tests for VTE do not perform as well in the heart failure population. D-dimer levels are already elevated in the majority of CHF patients; thus, its utility for suspected VTE in this population is severely compromised. Lung scintigraphy is often non-diagnostic, but computed tomography maintains its accuracy and is the preferred imaging test to evaluate suspected PE in CHF patients. The therapeutic options in CHF patients with VTE are the same and include the commonly used anticoagulants, fibrinolytics, and catheter or surgical embolectomy. Nonetheless, comorbidities, numerous medications, and renal dysfunction—all commonly seen in CHF patients—pose serious challenges to the management of patients with PE.

In summary, CHF is an important risk factor for VTE, with an incidence that varies widely from <1% to 20% depending on the severity of the disease and the clinical context. Both conditions negatively affect each other's prognosis, and therefore a high index of suspicion should be maintained in patients presenting with decompensated CHF. The preferred method for confirming and excluding the diagnosis of PE in the CHF population is the computed tomography scan. Treatment should be promptly initiated once the diagnosis is suspected, and careful attention should be paid to the choice, dosing, and management of antithrombotic therapy, as several physiologic abnormalities place these patients at a particularly high risk of bleeding complications.

Anticoagulation in atrial fibrillation: an important issue

Atrial fibrillation strokes are associated with a 30-day mortality of 24%. Warfarin has been shown to reduce the risk of stroke by 66%. Despite its proven efficacy, warfarin is underused in clinical practice. The dose response of warfarin is affected by age, sex, weight, liver function, dietary vitamin K, drugs, and pharmacogenomic factors. The narrow therapeutic window, coupled with a highly variable dose response, mandates frequent monitoring of the INR, which poses a barrier to warfarin's effectiveness in clinical practice. Several alternatives to warfarin have been evaluated in clinical trials. The first, dabigatran, is a direct oral thrombin inhibitor. Both rivaroxaban and apixaban inhibit factor Xa. All three of these agents reduced the risk of stroke (composite end point of ischemic and

hemorrhagic stroke) among patients with AF. The reduction in intracranial hemorrhage is unprecedented and is one of the most remarkable features of these drugs. Apixaban and the lower dose of dabigatran (110 mg) also reduced major extracranial bleeding compared with warfarin. These new oral anticoagulants are characterized by shorter half-lives compared with warfarin, do not require monitoring, and have fewer drug interactions. The degree of renal clearance is an important distinguishing feature of these drugs, as is the dosing frequency. Approximately 80% of dabigatran is eliminated by the kidneys, 66% of rivaroxaban (36% as unchanged drug), and 25% of apixaban. Both dabigatran and apixaban are dosed twice daily, and rivaroxaban is taken once daily.

Translating the efficacy of the novel anticoagulants from randomized trials into clinical practice will require heightened vigilance around medication adherence and changes in renal function. Creatinine clearance needs to be measured prior to initiation of these agents and then periodically throughout the duration of therapy. Although drug interactions are significantly less common with the new anticoagulants compared with warfarin, all are substrates of the P-glycoprotein (P-gp) transporter, so the potential for interaction with P-gp inhibitors and P-gp inducers exists. In addition, both apixaban and rivaroxaban are metabolized via CYP3A4. The short half-life and rapid onset of action obviate the need for perioperative bridging in most instances, but also highlight the importance of hemostasis prior to initiation following invasive procedures. The safety of these agents in combination with dual antiplatelet therapy warrants further study. In addition, data are needed regarding reversal of these anticoagulants in the setting of trauma, urgent surgery, and major hemorrhage. The ability to monitor the anticoagulant effect in select clinical situations remains a priority for real-world practice. The new oral anticoagulants represent a major advance in the prevention and treatment of thromboembolic disease. Current and planned studies will continue to inform their optimal use.

Stroke and bleeding in atrial fibrillation: how should we assess them?

The prevalence of AF in the United States is expanding rapidly. By 2050, estimates project that there will be over 5 million Americans with AF. Patients with both paroxysmal and persistent AF face elevated risks for stroke and systemic emboli. This risk can be successfully mitigated if patients are treated with chronic anticoagulation, yet use of such therapy can also cause serious bleeding events including intracranial hemorrhage and even death. Thus, selection of anticoagulant therapy should ideally be individualized, guided by a patient's specific risks for stroke with AF and bleeding with anticoagulation.

These risks can now be accurately estimated used several published risk prediction models. The $CHADS_2$ risk score uses a simple additive sum of five clinical features (CHF, hypertension, age $\geq$ 75 years, diabetes, and stroke [weighted = 2]) to stratify risk for stroke in AF. The $CHAD_2$-VASc score extends this simple risk score by adding in three additional risk factors: vascular disease, age 65–75 years, and sex = female). This modification provides for slightly more accurate risk estimation, particularly among lower-risks groups. Clinical trials have demonstrated that the benefits of warfarin therapy are linearly associated with patient risks. Based on these data, current U.S. and European guidelines recommend initiation of anticoagulation in all patients with moderate-to-high stroke risk.

The risks of bleeding with warfarin therapy are also associated with certain patient-related risk factors, which can be summated using published risk scores. These include the HAS-BLED, ATRIA, and HEMORR2HAGES bleeding risk scores. While these scores can stratify risk, the scores are often based on retrospective factors (such as labile INR results), thereby limiting their use in prospectively selecting a given patient for drug initiation. These bleeding scores also unfortunately demonstrate that many patient factors associated with stroke risk (including age, hypertension, and stroke) are also risk factors for bleeding. For example, the choices for anticoagulation therapy in older patients with multiple comorbidities will require a trade-off of high benefit and high risks.

The individualization of therapy does not end with these risk factors but is also determined, in part, by provider and system factors. For example, concomitant medications can also raise a patient's risk for bleeding. Specifically, combining use of aspirin and/or an ADP-inhibitor with warfarin can substantially raise a patient's risk for bleeding. Additionally, the effective and safe therapeutic window for warfarin therapy is narrow. Therefore, close patient monitoring and appropriate dose titration are needed to achieve optimal results. However, data indicate that this is rarely achieved in current U.S. clinical practice, where patients on warfarin are within the recommended therapeutic range less than 55% of the time. Newer anticoagulant therapies, including the direct thrombin inhibitor (dabigatran) and the new factor Xa inhibitors (apixaban and rivaroxaban), offer the promise of reducing anticoagulation-related bleeding risks, particularly the dreaded risk of intracranial hemorrhage. Yet, while these novel drugs may not require such careful therapeutic titration, patients must still closely and continuously adhere to instructions for the drugs to be safe and effective. These later points emphasize the importance of considering patient, provider, and system factors when

selecting AF treatment, as well as highlight future opportunities for quality improvement in AF care.

Triple therapy

Overlapping indications for antithrombotic therapy may lead to the need for "triple therapy," defined currently as aspirin, clopidogrel, and oral anticoagulation.

As the population ages, more patients will have both ACS and AF; accordingly triple therapy may be used more frequently. Prior studies have shown that, with more antithrombotic therapy, risk of bleeding increases. Many antiplatelet and anticoagulant drugs are part of the foundation for treatment of ACS and AF, making the decision about the right combination of these agents challenging. However, limited evidence is available to guide therapeutic decision-making about triple therapy. Registry information, subgroup analyses from clinical trials, and overviews of single-center experiences have been published, but no randomized trials evaluating different strategies of triple therapy have been completed.

Multiple guidelines and consensus statements from national societies provide recommendations for clinicians concerning the use of triple therapy. A simple flow diagram can be used by physicians to guide decisions about the need for dual antiplatelet therapy or triple therapy based on the assessment of patient bleeding and stroke risk. Five additional factors should be considered: (1) use of the lowest dose of antiplatelet therapy; (2) use of bare metal stents versus drug-eluting stents to minimize the duration of antiplatelet therapy; (3) optimal INR within a range of 2.0–2.5; (4) gastric protection with proton-pump inhibitors; and (5) minimization of the duration of triple therapy. It is also important to re-evaluate regularly the need for triple therapy. The risk of stent thrombosis will decrease over time, whereas bleeding risk will remain constant.

The ongoing What is the Optimal Antiplatelet and Anticoagulant Therapy in Patients with Oral Anticoagulation and Coronary Stenting (WOEST) trial is a prospective, multicenter, open-label randomized trial that aims to determine whether the combination of oral anticoagulants and clopidogrel 75 mg/d reduces the risk of bleeding and is not inferior to triple therapy (clopidogrel + oral anticoagulants + aspirin) with respect to the prevention of thrombotic complications. The primary outcome of the study will be the occurrence of bleeding up to 30 days and one year. Major adverse cardiac events will be the secondary outcomes. Sample size is 496, and the estimated study completion date is October 2011 (www.clinicaltrials.gov, NCT00769938).

The Triple Therapy in Patients on Oral Anticoagulation After Drug-Eluting Stent Implantation (ISAR-TRIPLE) study, an interventional, randomized, open-label trial, was designed to compare the six-week versus six-month clopidogrel treatment regimen in patients with concomitant aspirin and oral anticoagulants following drug-eluting stenting. The composite of death, MI, definite stent thrombosis, stroke, or major bleeding will be the primary outcome. The secondary outcomes will be the composite of cardiac death, MI, stent thrombosis, or ischemic stroke, as well as major bleeding complications. Estimated enrollment is 600, and the completion date is July 2012 (www.clinicaltrials.gov, NCT00776633). These two studies will also provide important insights about the use of triple therapy in clinical practice.

The optimal antithrombotic strategy for patients with ACS and AF who do or do not undergo PCI is still uncertain. Based on the available data, triple therapy offers the best protection against stroke and myocardial events but at the cost of increased bleeding. Currently, triple therapy defined as aspirin plus clopidogrel plus warfarin should be administered for the shortest period possible. Triple therapy may be redefined in the near future with new $P2Y_{12}$ inhibitors such as prasugrel and ticagrelor, newer antiplatelet agents such as PAR-1 inhibitors, other oral factor Xa inhibitors such as rivaroxaban or apixaban, and antithrombin agents such as dabigatran.

Recent data suggest that some new antithrombotic agents under development have significant potential to improve anticoagulant therapy. Their uptake and use will depend mostly on efficacy, safety, and cost relative to current medications. Moreover, a careful balance of antithrombotic efficacy and bleeding risk is now recognized as essential. Forecasting the death of vitamin K antagonists such as warfarin may be still premature, however. How to apply therapies, even when they have been shown to provide benefits in trials, will continue to be a major challenge in clinical practice. Finally, the optimal duration of triple therapy use and the population that may benefit the most from it also need to be defined.

How to manage anticoagulation in AF patients undergoing cardioversion

The reversion of an abnormal heart rhythm to its normal state was first described by Bernard Lown in 1961. The initial experience with cardioversion of ventricular and supra ventricular arrhythmias, including atrial fibrillation and atrial flutter, was furthered by several clinicians practicing at prominent institutions, who documented successes and potential complications, among them systemic embolism and stroke.

Guidelines for the management of AF and atrial flutter to include the peri-cardioversion period have been

formulated by national and international organizations and underscore the importance of anticoagulant therapy to minimize the risk of cardioembolic events that can be life-threatening or life-altering. While either IV or oral anticoagulants can be used as thromboprophylaxis for patients undergoing cardioversion, there are two fundamental tenets of effective therapy. First, a sufficient intensity of anticoagulation must be achieved at the time of the procedure. Second, a threshold level of anticoagulation must be maintained during a vulnerable period of varied duration after successful restoration of normal sinus rhythm.

The limitations of bleeding risk prediction rules in clinical practice

Clinicians prescribe anticoagulant therapy when they decide that the risks of such therapy (major bleeding) are outweighed by its benefits (thrombosis prevention). Estimating the benefits of anticoagulation is relatively straightforward, especially in patients with AF, where the widely accepted $CHADS_2$ score correlates well with annual stroke risk. To help clinicians determine the trade-offs associated with anticoagulation, several groups have derived prediction models that use patient characteristics to stratify an individual's annual risk for major bleeding. Although the development of these bleeding risk prediction models has helped clinicians by identifying factors (e.g., concomitant antiplatelet therapy, anemia, renal/hepatic failure, history of stroke) that are independently associated with an increased risk of warfarin-associated major bleeding, the models themselves are of limited value in clinical practice for a number of reasons. First, the majority of patients with AF stand to gain so much benefit (by reducing the risk of ischemic stroke) from warfarin that almost no calculated risk of bleeding should preclude a trial of anticoagulation. But even for the minority of AF patients whose absolute stroke risk reduction from warfarin is low (e.g., $CHADS_2$ score = 1), where these bleeding risk calculators seem more applicable, their use is impractical. All bleeding risk models estimate the annual likelihood that a patient will experience major bleeding. Unlike AF-related stroke (which is fatal or severely disabling more than 50% of the time), the "major bleeding" predicted by these models represent clinical events within a spectrum that includes both simple blood transfusion as well as fatal intracranial hemorrhage. Because warfarin-related major bleeding is fatal in fewer than 10% if cases, it is difficult to know how to weigh the information generated by a bleeding prediction model against the benefit of reducing the risk of ischemic stroke.

Even if the trade-offs were more straightforward, several of these bleeding risk scores are, unlike the $CHADS_2$ score, difficult to remember and/or calculate. In the case of the HAS-BLED score, we must know whether the patient has "labile INR values"—this is information that cannot be known to the clinicians who are trying to decide whether to initiate warfarin treatment. In the case of the model published by Shireman, the equation used to define bleeding risk is so complicated that the authors admit clinicians would be unable to commit it to memory. Other models are incomplete: the ATRIA bleeding risk score does not account for concomitant antiplatelet therapy, a factor known to increase the risk of warfarin-associated major hemorrhage more than two-fold. Finally, none of the new oral anticoagulants being studied (or recently approved) for stroke prevention in AF has been evaluated by these bleeding prediction models.

In summary, there is no doubt that the benefits of anticoagulant therapy must be balanced against the risk that they might cause major bleeding. That notwithstanding, for the reasons outlined above, the currently available bleeding risk scores/models have limited utility in everyday clinical practice.

Thromboprophylaxis for medical patients: should it be the default?

Medical thromboprophylaxis reduces the risk of DVT, PE, and fatal PE. There is excellent quality evidence that medical prophylaxis is under-prescribed, resulting in otherwise avoidable episodes of VTE. Use of medical prophylaxis has been endorsed by numerous peer organizations and is the focus of initiatives such as Required Organizational Practices of Accreditation Canada (http://www.accreditation.ca/uploadedFiles/ROP%20Handbook.pdf, accessed November 14, 2011).

However, it is clear that not all medical patients require thromboprophylaxis. Some patients have contraindications to anticoagulant therapy (such as those who have active bleeding and those with severe acquired or congenital bleeding disorders). Mechanical forms of prophylaxis (such as intermittent pneumatic compression devices or graduated compression stockings) might be used in such patients; however, their efficacy has never been tested in well-performed prospective studies. Extrapolating from other situations, mechanical prophylaxis reduces the risk of VTE but is associated with the potential for toxicity, including reduced mobilization, spread of infection, and skin ulceration. Mechanical forms of prophylaxis are also poorly tolerated by most medical patients.

More importantly, and the focus of recent evidence, is the observation that many patients in the hospital are at very low risk of VTE. When prophylaxis is administered to these patients, it results in the potential for toxicity

(bleeding and heparin-induced thrombocytopenia), as well as expense with little likelihood of benefit (because the background rate of VTE is sufficiently low so that it is unlikely to be lowered further by therapy).

There is no doubt that all patients admitted to the hospital should receive consideration for the administration of anticoagulant prophylaxis or mechanical prophylaxis. Adherence to this recommendation will be an enhanced by the use of preprinted orders, computerized prompting, and a multidisciplinary approach to care wherein multiple individuals take responsibility for ensuring the provision of optimal prophylaxis. When patients are considered to be at sufficiently low risk, then documentation of the rationale for non-administration of prophylaxis should be provided. This could take the form of a note in the chart or a specific notation in the computerized order entry system discussing why prophylaxis is not being administered.

Various authors have attempted to provide guidance for clinicians with respect to patient selection for prophylaxis. In a widely referenced paper, Dr. Charles Francis has proposed that a combination of age more than 40 years, anticipated limited mobility of three or more days, and at least one other risk factor provide sufficient likelihood of VTE to warrant prophylaxis. Conditions associated with increased risk include, but are not limited to, CHF, MI, stroke, inflammatory bowel disease, and a history of previous VTE, recent surgery, trauma or immobilization, obesity, central venous catheterization, known acquired or inherited thrombophilic states, and, potentially, estrogen therapy. Patients meeting three or more criteria should probably receive prophylaxis. In many hospitals, this will encompass nearly all patients; however, there remains a small subset of patients who do not meet criteria and who probably have a sufficiently low risk of VTE that the use of antithrombotic prophylaxis cannot be justified.

The case against routine prophylaxis was probably most compellingly made by Barbar and colleagues in a recently published paper. One thousand one hundred and eighty consecutive patients were followed after admission to an internal medicine service for the development of VTE over 90 days. A previously validated scoring system was used to assign patients to various risk classes for the development of VTE. Fully 711 patients were found to have a low risk of venous thromboembolism, of whom 659 (93%) did not receive any VTE prophylaxis. Among the 711 patients, there were two clinically apparent episodes of venous thrombi embolism, one occurring in the 659 patients who did not receive prophylaxis and one occurring in the 52 patients who did. These observations argue compellingly against the use of routine prophylaxis because the administration of routine heparin or LMWH to the 659 patients would have exposed them to the potential for toxicity without hope of further reducing the very low risk of VTE observed in this study. There is absolutely no doubt that bleeding rates would "swamp" the potential benefit of a reduction in the rate of VTE.

In summary, the administration of pharmacologic or mechanical thromboembolism prophylaxis is a critical consideration for all patients admitted to the hospital. Many patients harbor significant risks for venous thromboembolism, and all such patient should be treated with effective forms of pharmacologic prophylaxis. However, a significant proportion of patients will have a sufficiently low risk of VTE that the risks and costs of pharmacologic and mechanical prophylaxis cannot be justified; in such patients, documentation of the rationale for the omission of therapy is required, but venous thromboembolism prophylaxis should consist simply of aggressive mobilization and early discharge from the hospital.

Acknowledgments The Fourth International Symposium of Thrombosis and Anticoagulation (ISTA) was sponsored by unrestricted grants from AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, and Daiichi Sankyo. The symposium was organized by the Brazilian Clinical Research Institute, and the chairmen of the meeting thank Cecilia Jahnel, Fabio Villas Boas, and Jadelson Andrade for all of their assistance in organizing the event.

Recommended reading

Ageno W, Garcia D et al (2009) Prevention and treatment of bleeding complications in patients receiving vitamin K antagonists, part 2: treatment. Am J Hematol 84(9):584–588

Agnelli G, Bergqvist D et al, On Behalf of the PEGASUS investigators (2005) Randomized clinical trial of postoperative fondaparinux versus perioperative dalteparin for prevention of venous thromboembolism in high-risk abdominal surgery. Br J Surg 92:1212–1222

Alexander JH, Lopes RD et al (2011) Apixaban with antiplatelet therapy after acute coronary syndrome. N Engl J Med 365(8): 699–708

Alexander JH, Becker RC et al (2009) Apixaban, an oral, direct, selective factor Xa inhibitor, in combination with antiplatelet therapy after acute coronary syndrome: results of the Apixaban for Prevention of Acute Ischemic and Safety Events (APPRAISE) trial. Circulation 119(22):2877–2885

Alexander KP, Peterson ED (2010) Minimizing the risks of anticoagulants and platelet inhibitors. Circulation 121:1960–1970

Anderson FA Jr, Zayaruzny M et al (2007) Estimated annual numbers of US acute-care hospital patients at risk for venous thromboembolism. Am J Hematol 82(9):777–782

Anderson WG, Alexander SC et al (2008) 'What concerns me is…' Expression of emotion by advanced cancer patients during outpatient visits. Support Care Cancer 16(7):803–811

Angiolillo DJ, Gibson CM et al (2011) Differential effects of omeprazole and pantoprazole on the pharmacodynamics and pharmacokinetics of clopidogrel in healthy subjects: randomized, placebo-controlled, crossover comparison studies. Clin Pharmacol Ther 89(1):65–74

Angiolillo DJ, Shoemaker SB et al (2007) Randomized comparison of a high clopidogrel maintenance dose in patients with diabetes mellitus and coronary artery disease: results of the Optimizing

Antiplatelet Therapy in Diabetes Mellitus (OPTIMUS) study. Circulation 115(6):708–716

Ansell J, Hirsh J et al (2008) Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians evidence-based clinical practice guidelines (8th edition). Chest 133(6 Suppl):160S–198S

Antiplatelet Trialists Collaboration (1994) Collaborative overview of randomised trials of antiplatelet therapy—III: reduction in venous thrombosis and pulmonary embolism by antiplatelet prophylaxis among surgical and medical patients. BMJ 308: 235–246

Atrial Fibrillation Investigators (1994) Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. Arch Intern Med 154:1449–1457

Ay C, Dunkler D et al (2010) Prediction of venous thromboembolism in cancer patients. Blood 116(24):5377–5382

Baker WL, Cios DA et al (2009) Meta-analysis to assess the quality of warfarin control in atrial fibrillation patients in the United States. J Manag Care Pharm 15:244–252

Barbar S, Noventa F et al (2010) A risk assessment model for the identification of hospitalized medical patients at risk for venous thromboembolism: the Padua Prediction Score. J Thromb Haemost 8(11):2450–2457

Berger JS, Frye CB et al (2008) Impact of clopidogrel in patients with acute coronary syndromes requiring coronary artery bypass surgery: a multicenter analysis. J Am Coll Cardiol 52(21): 1693–1701

Beyth RJ, Quinn LM et al (1998) Prospective evaluation of an index for predicting the risk of major bleeding in outpatients treated with warfarin. Am J Med 105:91–99

Bhatt DL, Cryer BL et al (2010) Clopidogrel with or without omeprazole in coronary artery disease. N Engl J Med 363(20): 1909–1917

Brandt JT, Close SL et al (2007) Common polymorphisms of CYP2C19 and CYP2C9 affect the pharmacokinetic and pharmacodynamic response to clopidogrel but not prasugrel. J Thromb Haemost 5(12):2429–2436

Breet NJ, van Werkum JW et al (2010) Comparison of platelet function tests in predicting clinical outcome in patients undergoing coronary stent implantation. JAMA 303(8):754–762

Camm AJ, Kirchhof P et al (2010) Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology. Eur Heart J 31:2369–2429

Campo G, Parrinello G et al (2011) Prospective evaluation of on-clopidogrel platelet reactivity over time in patients treated with percutaneous coronary intervention relationship with gene polymorphisms and clinical outcome. J Am Coll Cardiol 57(25): 2474–2483

Caprini JA (2010) Mechanical methods for thrombosis prophylaxis. Clin Appl Thromb Hemost 16(6):668–673

Chan MY, Cohen MG et al (2008) Phase 1b randomized study of antidote-controlled modulation of factor IXa activity in patients with stable coronary artery disease. Circulation 117(22): 2865–2874

Cohen AT, Tapson VF et al (2008) Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): a multinational cross-sectional study. Lancet 371(9610): 387–394

Connolly SJ, Ezekowitz MD et al (2009) Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 361(12): 1139–1151

Cook D, Meade M et al (2011) Prophylaxis for thromboembolism in critical care trial protocol and analysis plan. J Crit Care 26(223):e1–9

Crowther MA, Warkentin TE (2009) Managing bleeding in anticoagulated patients with a focus on novel therapeutic agents. J Thromb Haemost 7(Suppl 1):107–110

CURRENT-OASIS 7 Investigators (2010) Dose comparisons of clopidogrel and aspirin in acute coronary syndromes. N Engl J Med 363:930–942

Eikelboom JW, Weitz JI (2008) Selective factor Xa inhibition for thromboprophylaxis. Lancet 372(9632):6–8

Eikelboom JW, Mehta SR et al (2006) Adverse impact of bleeding on prognosis in patients with acute coronary syndromes. Circulation 114(8):774–782

Eriksson BI, Borris LC et al (2008) Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty. N Engl J Med 358(26):2765–2775

Eriksson BI, Dahl OE et al (2007) Oral dabigatran etexilate versus subcutaneous enoxaparin for the prevention of venous thromboembolism after total knee replacement: the RE-MODEL randomized trial. J Thromb Haemost 5(11):2178–2185

Fang MC, Go AS et al (2011) A new risk scheme to predict warfarin-associated hemorrhage: the ATRIA (Anticoagulation and Risk Factors in Atrial Fibrillation) study. J Am Coll Cardiol 58:395–401

Ferraris VA, Brown JR et al (2011) 2011 update to the Society of Thoracic Surgeons and Society of Cardiovascular Anesthesiologists blood conservation clinical practice guidelines. Ann Thorac Surg 91:944–982

Fitchett D, Eikelboom J et al (2009) Dual antiplatelet therapy in patients requiring urgent coronary artery bypass grafting surgery: a position statement of the Canadian Cardiovascular Society. Can J Cardiol 25(12):683–689

Francis CW (2007) Clinical practice. Prophylaxis for thromboembolism in hospitalized medical patients. N Engl J Med 356(14): 1438–1444

Gage BF, Yan Y et al (2006) Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). Am Heart J 151:713–719

Gage BF, Waterman AD et al (2001) Validation of clinical classification schemes for predicting stroke. JAMA 285: 2864–2870

Garcia D, Libby E et al (2010) The new oral anticoagulants. Blood 115(1):15–20

Geerts WH, Bergqvist D et al (2008) Prevention of venous thromboembolism: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 133(6 Suppl):381S–453S

Gharacholou SM, Alexander KP et al (2010) Implications and reasons for the lack of use of reperfusion therapy in patients with ST-segment elevation myocardial infarction: findings from the CRUSADE initiative. Am Heart J 159(5):757–763

Gibson CM, Mega JL et al (2011) Rationale and design of the Anti-Xa therapy to lower cardiovascular events in addition to standard therapy in subjects with acute coronary syndrome–thrombolysis in myocardial infarction 51 (ATLAS-ACS 2 TIMI 51) trial: a randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of rivaroxaban in subjects with acute coronary syndrome. Am Heart J 161(5):815–821

Gilard M, Arnaud B et al (2008) Influence of omeprazole on the antiplatelet action of clopidogrel associated with aspirin: the randomized, double-blind OCLA (Omeprazole CLopidogrel Aspirin) study. J Am Coll Cardiol 51(3):256–260

Goodman SG, Clare R et al (2012) Association of proton-pump inhibitor use on cardiovascular outcomes with clopidogrel and ticagrelor: insights from PLATO. Circulation Jan 18. [Epub ahead of print]

Granger CB, Alexander JH et al (2011) Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med 365:981–992

Gurbel PA, Bliden KP et al (2009) Randomized double-blind assessment of the ONSET and OFFSET of antiplatelet effects of ticagrelor versus clopidogrel in patients with stable coronary artery disease: the ONSET/OFFSET study. Circulation 120: 2577–2585

Gurbel PA, Becker RC et al (2007) Platelet function monitoring in patients with coronary artery disease. J Am Coll Cardiol 50(19): 1822–1834

Hamm CW, Bassand J-P et al (2011) ESC guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: the task force for the management of acute coronary syndromes (ACS) in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J [Epub ahead of print]

Hansen ML, Sorensen R et al (2010) Risk of bleeding with single, dual, or triple therapy with warfarin, aspirin, and clopidogrel in patients with atrial fibrillation. Arch Intern Med 170(16): 1433–1441

Harrington RA, Becker RC et al (2008) Antithrombotic therapy for non-ST-segment elevation acute coronary syndromes: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 133(6 Suppl):670S–707S

Heit JA (2008) The epidemiology of venous thromboembolism in the community. Arterioscler Thromb Vasc Biol 28(3):370–372

Held C, Åsenblad N et al (2011) Ticagrelor versus clopidogrel in patients with acute coronary syndromes undergoing coronary artery bypass graft surgery: results from the PLATO (Platelet Inhibition and Patient Outcomes) trial. J Am Coll Cardiol 57:672–684

Hillis LD, Smith PK et al (2011) 2011 ACCF/AHA guideline for coronary artery bypass graft surgery. J Am Coll Cardiol [Published online November 7, 2011]

Hirsh J, Guyatt G et al (2008a) Antithrombotic and thrombolytic therapy: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 133(6):110S–112S

Hirsh J, Guyatt G et al (2008b) Executive summary: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 133(6 Suppl):71S–109S

Hochholzer W, Trenk D et al (2010) Impact of cytochrome P450 2C19 loss-of-function polymorphism and of major demographic characteristics on residual platelet function after loading and maintenance treatment with clopidogrel in patients undergoing elective coronary stent placement. J Am Coll Cardiol 55:2427–2434

Hochholzer W, Trenk D et al (2005) Time dependence of platelet inhibition after a 600-mg loading dose of clopidogrel in a large, unselected cohort of candidates for percutaneous coronary intervention. Circulation 111(20):2560–2564

Kakkar AK, Brenner B et al (2008) Extended duration rivaroxaban versus short-term enoxaparin for the prevention of venous thromboembolism after total hip arthroplasty: a double-blind, randomised controlled trial. Lancet 372(9632):31–39

Karthikeyan G, Eikelboom JW et al (2009) Does acetyl salicylic acid (ASA) have a role in the prevention of venous thromboembolism? Br J Haematol 146:142–149

Kearon C, O'Donnell M (2011) Graduated compression stockings to prevent venous thromboembolism in hospital: evidence from patients with acute stroke. Pol Arch Intern Med 121(1–2):40–43

Keegan SP, Patrick DM et al (2009) Prevention of perioperative venous thromboembolism. Int Anesthesiol Clin 47(4):55–64

Khorana AA, Kuderer NM et al (2008) Development and validation of a predictive model for chemotherapy-associated thrombosis. Blood 111(10):4902–4907

Kushner FG, Hand M et al (2009) 2009 Focused updates: ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction (updating the 2004 guideline and 2007 focused update) and ACC/AHA/SCAI guidelines on percutaneous coronary intervention (updating the 2005 guideline and 2007 focused update) a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 54(23):2205–2241

Lachiewicz PF (2009) Comparison of ACCP and AAOS guidelines for VTE prophylaxis after total hip and total knee arthroplasty. Orthopedics 32:74–78

Lassen MR, Raskob GE et al (2009) Apixaban or enoxaparin for thromboprophylaxis after knee replacement. N Engl J Med 361(6):594–604

Lassen MR, Ageno W et al (2008) Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty. N Engl J Med 358(26):2776–2786

Leiria TLL, Lopes RD et al (2011) Antithrombotic therapies in patients with prosthetic heart valves: guidelines translated for the clinician. J Thromb Thrombolysis 31:514–522

Leissinger CA, Blatt PM et al (2008) Role of prothrombin complex concentrates in reversing warfarin anticoagulation: a review of the literature. Am J Hematol 83(2):137–143

Lip GY, Nieuwlaat R et al (2010) Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 137(2):263–272

Lopes RD, Becker RC et al (2011) Highlights from the III international symposium of thrombosis and anticoagulation (ISTA), October 14–16, 2010, São Paulo, Brazil. J Thromb Thrombolysis 32(2):242–266

Lopes RD, Alexander KP (2009) Antiplatelet therapy in older adults with non–ST-segment elevation acute coronary syndrome: considering risks and benefits. Am J Cardiol 104(5 Suppl): 16C–21C

Lopes RD, Elliott LE et al (2009) Antithrombotic therapy and outcomes of patients with atrial fibrillation following primary percutaneous coronary intervention: results from the APEX-AMI trial. Eur Heart J 30(16):2019–2028

Lopes RD, Piccini JP et al (2008) Antithrombotic therapy in atrial fibrillation: guidelines translated for the clinician. J Thromb Thrombolysis 26(3):167–174

Lown B (1967) Electrical reversion of cardiac arrhythmias. Br Heart J 29:469–489

Mahaffey KW, Wojdyla DM et al, On Behalf of the PLATO Investigators (2011) Ticagrelor compared with clopidogrel by geographic region in the Platelet Inhibition and Patient Outcomes (PLATO) trial. Circulation 124:544–554

Matchar DB, Jacobson A et al (2010) Effect of home testing of international normalized ratio on clinical events. N Engl J Med 363(17):1608–1620

Mega JL, Close SL et al (2010) Genetic variants in ABCB1 and CYP2C19 and cardiovascular outcomes after treatment with clopidogrel and prasugrel in the TRITON–TIMI 38 trial: a pharmacogenetic analysis. Lancet 376:1312–1319

Mega JL, Braunwald E et al (2009) Rivaroxaban versus placebo in patients with acute coronary syndromes (ATLAS ACS-TIMI 46): a randomised, double-blind, phase II trial. Lancet 374(9683): 29–38

Mehta SR, Boden WE et al (2008) Antithrombotic therapy with fondaparinux in relation to interventional management strategy in patients with ST- and non-ST-segment elevation acute coronary syndromes: an individual patient-level combined analysis of the Fifth and Sixth Organization to Assess Strategies in Ischemic Syndromes (OASIS 5 and 6) randomized trials. Circulation 118(20):2038–2046

Montalescot G, Wiviot SD et al (2009) Prasugrel compared to clopidogrel in patients undergoing percutaneous coronary

intervention for ST-elevation myocardial infarction (TRITON-TIMI 38): double blind, randomized controlled trial. Lancet 373:723–731

Nagarakanti R, Ezekowitz MD et al (2011) Dabigatran versus warfarin in patients with atrial fibrillation: an analysis of patients undergoing cardioversion. Circulation 123(2):131–136

Nagarajan DV, Lewis PS et al (2004) Is addition of antiplatelet therapy to warfarin beneficial to patients with prosthetic heart valves? Interact Cardiovasc Thorac Surg 3:450–455

Nijjer SS, Watson G et al (2011) Safety of clopidogrel being continued until the time of coronary artery bypass surgery in acute coronary syndrome: a meta-analysis of 34 studies. Eur Heart J [Epub ahead of print]

O'Connell C, Razavi P et al (2011) Unsuspected pulmonary emboli adversely impact survival in patients with cancer undergoing routine staging multi-row detector computed tomography scanning. J Thromb Haemost 9(2):305–311

Oldgren J, Budaj A et al (2011) Dabigatran vs. placebo in patients with acute coronary syndromes on dual antiplatelet therapy: a randomized, double-blind, phase II trial. Eur Heart J [Epub ahead of print]

Oldgren J, Alings M et al (2010) Dabigatran versus warfarin in atrial fibrillation patients with low, moderate, and high $CHADS_2$ score: a RE-LY subgroup analysis. J Am Coll Cardiol 55(10A): A1.E2

Palumbo A, Rajkumar SV et al (2008) Prevention of thalidomide- and lenalidomide-associated thrombosis in myeloma. Leukemia 22(2):414–423

Patel MR, Mahaffey KW et al (2011) Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med 365:883–891

Pisters R, Lane DA et al (2010) A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. Chest 138:1093–1100

Price MJ, Berger PB et al, For the GRAVITAS Investigators (2011) Standard-vs. high-dose clopidogrel based on platelet function testing after percutaneous coronary intervention: the GRAVITAS randomized trial. JAMA 305:1097–1105

Price MJ, Endemann S et al (2008) Prognostic significance of post-clopidogrel platelet reactivity assessed by a point-of-care assay on thrombotic events after drug-eluting stent implantation. Eur Heart J 29(8):992–1000

PROTECT Investigators for the Canadian Critical Care Trials Group and the Australian and New Zealand Intensive Care Society Clinical Trials Group, Cook D, Meade M et al (2011) Dalteparin versus unfractionated heparin in critically ill patients. N Engl J Med 364:1305–1314

Rodgers A, MacMahon S et al (2000) Prevention of pulmonary embolism and deep vein thrombosis with low dose aspirin: Pulmonary Embolism Prevention (PEP) trial. Lancet 355: 1295–1302

Rothberg MB, Celestin C et al (2005) Warfarin plus aspirin after myocardial infarction or the acute coronary syndrome: meta-analysis with estimates of risk and benefit. Ann Intern Med 143(4):241–250

Santos AT, Kalil RA et al (2006) A randomized, double-blind, and placebo-controlled study with tranexamic acid of bleeding and fibrinolytic activity after primary coronary artery bypass grafting. Braz J Med Biol Res 39(1):63–69

Sibbing D, Stegherr J et al (2010) A double-blind, randomized study on prevention and existence of a rebound phenomenon of platelets after cessation of clopidogrel treatment. J Am Coll Cardiol 55(6):558–565

Sibbing D, Braun S et al (2009) Platelet reactivity after clopidogrel treatment assessed with point-of-care analysis and early drug-eluting stent thrombosis. J Am Coll Cardiol 53(10):849–856

Siller-Matula JM, Jilma B et al (2010) Effect of proton pump inhibitors on clinical outcome in patients treated with clopidogrel: a systematic review and meta-analysis. J Thromb Haemost 8(12):2624–2641

Singer DE, Chang Y et al (2009) The net clinical benefit of warfarin anticoagulation in atrial fibrillation. Ann Intern Med 151:297–305

Singer DE, Albers GW et al (2008) Antithrombotic therapy in atrial fibrillation: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 133(6 Suppl):546S–592S

Steg PG, Mehta SR et al (2011) RUBY-1: a randomized, double-blind, placebo-controlled trial of the safety and tolerability of the novel oral factor Xa inhibitor darexaban (YM150) following acute coronary syndrome. Eur Heart J 32(20):25–41

Steg PG, FitzGerald G et al (2009) Risk stratification in non–ST-segment elevation acute coronary syndromes: troponin alone is not enough. Am J Med 122(2):107–108

Subherwal S, Bach RG et al (2009) Baseline risk of major bleeding in non-ST-segment-elevation myocardial infarction: the CRUSADE (Can Rapid risk stratification of Unstable angina patients Suppress ADverse outcomes with Early implementation of the ACC/AHA Guidelines) Bleeding Score. Circulation 119(14): 1873–1882

Thygesen K, Alpert JS et al (2007) Universal definition of myocardial infarction. J Am Coll Cardiol 50(22):2173–2195

van Giezen JJ, Sidaway J et al (2011) Ticagrelor inhibits adenosine uptake in vitro and enhances adenosine-mediated hyperemia responses in a canine model. J Cardiovasc Pharmacol Ther [Epub ahead of print]

Waldo AL, Becker RC et al, NABOR Steering Committee (2005) Hospitalized patients with atrial fibrillation and a high risk of stroke are not being provided with adequate anticoagulation. J Am Coll Cardiol 46(9):1729–1736

Wallentin L, Yusuf S et al (2010) Efficacy and safety of dabigatran compared with warfarin at different levels of international normalised ratio control for stroke prevention in atrial fibrillation: an analysis of the RE-LY trial. Lancet 376(9745):975–983

Wallentin L (2009) P2Y(12) inhibitors: differences in properties and mechanisms of action and potential consequences for clinical use. Eur Heart J 30(16):1964–1977

Wallentin L, Becker RC et al (2009) Ticagrelor versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 361(11):1045–1057

Wann LS, Curtis AB et al (2011) 2011 ACCF/AHA/HRS focused update on the management of patients with atrial fibrillation (update on dabigatran). A report of the American College of Cardiology Foundation Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 57(11): 1330–1337

Wann LS, Curtis AB et al (2011) 2011 ACCF/AHA/HRS focused update on the management of patients with atrial fibrillation (updating the 2006 guideline). J Am Coll Cardiol 57:223–242

Wiviott SD, Braunwald E et al (2008) Greater clinical benefit of more intensive oral antiplatelet therapy with prasugrel in patients with diabetes mellitus in the trial to assess improvement in therapeutic outcomes by optimizing platelet inhibition with prasugrel-Thrombolysis in Myocardial Infarction 38. Circulation 118(16): 1626–1636

Wiviott SD, Braunwald E et al (2007) Prasugrel versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 357(20):2001–2015

Wright RS, Anderson JL et al (2011) 2011 ACCF/AHA focused update of the guidelines for the management of patients with unstable angina/non–ST-elevation myocardial infarction (updating the

2007 guideline): a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation 123:2022–2060

Yang X, Alexander KP et al (2005) The implications of blood transfusions for patients with non-ST-segment elevation acute coronary syndromes: results from the CRUSADE National Quality Improvement Initiative. J Am Coll Cardiol 46(8):1490–1495

Yusuf S, Mehta SR et al (2006a) Comparison of fondaparinux and enoxaparin in acute coronary syndromes. N Engl J Med 354(14): 1464–1476

Yusuf S, Mehta SR et al (2006b) Effects of fondaparinux on mortality and reinfarction in patients with acute ST-segment elevation myocardial infarction: the OASIS-6 randomized trial. JAMA 295(13):1519–1530

Journal of Perinatology (2012) 32, 317–324

www.nature.com/jp

STATE-OF-THE-ART

Thermal protection of the newborn in resource-limited environments

K Lunze[1] and DH Hamer[2,3,4,5]

[1]Preventive Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA; [2]Center for Global Health and Development, Boston University, Boston, MA, USA; [3]Department of International Health, Boston University School of Public Health, Boston, MA, USA; [4]Section of Infectious Diseases, Department of Medicine, Boston University School of Medicine, Boston, MA, USA and [5]Zambia Centre for Applied Health Research and Development, Lusaka, Zambia

Appropriate thermal protection of the newborn prevents hypothermia and its associated burden of morbidity and mortality. Yet, current global birth practices tend to not adequately address this challenge. Here, we discuss the pathophysiology of hypothermia in the newborn, its prevention and therapeutic options with particular attention to resource-limited environments. Newborns are equipped with sophisticated mechanisms of body temperature regulation. Neonatal thermoregulation is a critical function for newborn survival, regulated in the hypothalamus and mediated by endocrine pathways. Hypothermia activates cellular metabolism through shivering and non-shivering thermogenesis. In newborns, optimal temperature ranges are narrow and thermoregulatory mechanisms easily overwhelmed, particularly in premature and low-birth weight infants. Hyperthermia most commonly is associated with dehydration and potentially sepsis. The lack of thermal protection promptly leads to hypothermia, which is associated with detrimental metabolic and other pathophysiological processes. Simple thermal protection strategies are feasible at community and institutional levels in resource-limited environments. Appropriate interventions include skin-to-skin care, breastfeeding and protective clothing or devices. Due to poor provider training and limited awareness of the problem, appropriate thermal care of the newborn is often neglected in many settings. Education and appropriate devices might foster improved hypothermia management through mothers, birth attendants and health care workers. Integration of relatively simple thermal protection interventions into existing mother and child health programs can effectively prevent newborn hypothermia even in resource-limited environments.
Journal of Perinatology (2012) **32,** 317–324; doi:10.1038/jp.2012.11; published online 1 March 2012

Keywords: newborn; neonatal; hypothermia; thermal protection

Introduction

The need for thermal newborn protection has long been known, as alluded to by Soranus of Ephesus (98 to 138 AD) in his four-volume treatise 'On Diseases on Women', which demonstrates the importance of keeping newborns warm.[1] The Bible provides probably the most well-known example of thermal protection of the newborn in Luke 2: 7, 'And she brought forth her firstborn son, and wrapped him in swaddling clothes, and laid him in a manger.'

Yet, in current times, the majority of the almost 4 million newborns globally who do not survive their first month of life[2] die of complications associated with hypothermia, such as prematurity and severe infections (mostly sepsis and pneumonia).[3] As two recent reviews have acknowledged, neonatal deaths related to hypothermia are relatively neglected, but considered easily preventable with attention to warmth, feeding and infection management.[4,5] This article focuses on the diagnosis of hypothermia and management of thermal protection of newborns in low-resource environments. We review mechanisms of neonatal thermoregulation, discuss the pathophysiology of newborn hypothermia and present simple strategies of thermal protection for the newborn.

Neonatal thermoregulation

The normal body temperature of a newborn infant is usually defined as ranging between 36.5 and 37.5 °C (97.7 to 99.5 °F).[6] A series of observational randomized trials starting in the late fifties[7,8] showed that keeping babies warm reduces mortality and morbidity, and spurred further research on the pathophysiology of thermoregulation in newborns. Thermoregulation is a biological priority for all homeothermic species.[9] Newborns, particularly preterm and low-birth weight (LBW) infants, have limited capacity for thermoregulation during the first weeks of life. The optimal environmental temperature is termed thermal neutral temperature, at which metabolic requirements of the organism are minimal.[10] Both a decreased and an increased core temperature increase the metabolic rate of newborns,[11] who have only very limited ability to maintain a normal temperature and easily become hypothermic or hyperthermic. Although hyperthermia also increases energy needs, hypothermia seems to carry a higher risk of complications.[12]

Correspondence: Dr K Lunze, Preventive Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA.
E-mail: karsten.lunze@post.harvard.edu
Received 28 September 2011; revised 21 December 2011; accepted 24 January 2012; published online 1 March 2012

When the infant's body temperature decreases in response to sudden exposure to cold extrauterine environments, signals from peripheral and central thermoreceptors reach the hypothalamus through afferent pathways.[13] The resulting norepinephrine release then triggers nonshivering thermogenesis, or lipolysis of brown adipose tissue, which is the main homeothermic heat production mechanism in newborns. Heat production occurs through uncoupling ATP synthesis via the oxidation of fatty acids in the mitochondria, utilizing uncoupled protein.[14]

Afferent temperature information is processed in the hypothalamus. Thermoregulation requires an intact central nervous system,[15] and impaired thermoregulation, either hypo- or hyperthermia, can be indicators of central nervous system damage. The hypothalamus has a central role in regulating the autonomic, somatic and endocrine systems to maintain a normal body temperature. Decreasing body temperatures trigger a release of thyroid-stimulating hormone, which leads to an increase in thyroxine and consequently triiodothyronine. The resulting norepinephrine release causes vasoconstriction, glycolysis and uncoupling of mitochondrial oxidation in the brown adipose tissue, further generating heat production.[14] The latter process is ineffective in preterm infants, because it depends on the amount of brown fat as well as levels of the enzymes 5′/3′-monodeiodinase and thermogenin, which build up only later in fetal development.[13]

Shivering is not regularly involved in a newborn's reaction to cold stress.[16] Another mechanism of heat production is infant behavior:[10] the irritable baby prompts the mother to hold the baby, drying, cuddling and swaddling him or her, thus preventing heat loss.

Newborns are unable to maintain their body temperature on their own without thermal protection. Although a newborn's thermoregulation is as complex as in adults if not more sophisticated, as discussed above, their range of optimal or even tolerable body temperature is narrower. A newborn placed naked in an environment of 23 °C at birth suffers the same cold as does a naked adult at 0 °C.[17] Without thermal protection, human neonates are functionally poikilothermic, that is, they change their body temperature according to environmental temperatures. In newborns placed in a colder environment, core temperature decreases at a rate 0.2 to 1.0 °C per minute and finally may lead to death from cessation of metabolic activities.[10]

Pathophysiology of newborn hypothermia

The World Health Organization (WHO) defines neonatal hypothermia as a temperature below 36.5 °C (97.7 °F) and proposes the following classification:[17] Mild hypothermia, caused by cold stress, is classified as a body temperature range from 36 to 36.5 °C (96.8 to 97.7 °F) and is considered a cause for concern,[17] because the exposed infant begins to lose more heat than he or she can produce.[13] Moderate hypothermia is a body temperature from 32 to 36 °C (89.6 to 96.8 °F), indicating danger and requiring warming of the baby. According to the WHO classification, a body temperature of <32 °C (89.6 °F) is considered severe hypothermia, or cold injury, with a potentially grave outcome, and needs immediate skilled attention.

Heat loss occurs in several ways. The most common scenario is that of a wet baby who is not dried, and in whom evaporation of fluid from the skin leads to heat loss. Evaporation often occurs with amniotic fluids during the first minutes of life or with water after a baby is bathed. The energy loss is substantial: immediately at delivery, when the environmental temperature surrounding the baby drops from 37 °C in the maternal womb to the usually less warm air temperature, evaporative heat loss begins at a rate of 0.58 kcal ml^{-1} fluid evaporated.[10]

A baby placed naked on a cold surface loses heat through conduction. A newborn exposed to cool surrounding air or draughts will lose heat through convection. Radiation from cool objects next to the baby (for example, a cold wall) can also lower its body temperature. Unlike in adults, sweat secretion has little or no role in the thermoregulation of a newborn or preterm baby.[18]

As all data on hypothermia are from observational studies and prospective randomized trials without treating hypothermia are not permissive, the direction of causality for factors associated with hypothermia is not entirely clear.

Some argue that lowering body temperatures might increase metabolic processes to generate heat, which could lead to hypoglycemia and hypoxia in response to increasing energy demands.[13] As hypothermia and hypoglycemia both exacerbate hypoxia, this would reinforce a vicious circle[19] and could on one hand explain the mortality associated with hypothermia. However, studies have shown that hypothermia is not a risk factor for neonatal hypoglycemia in analyses adjusted for confounders such as LBW or anemia.[20] On the other hand, hypoglycemia is common among newborns in resource-limited settings[20] and, instead of being a consequence, could rather be a cause of hypothermia.

Similarly, with regard to the reported associations of hypothermia with infections and organ failure,[21,22] hypothermia might be either a consequence or a cause of severe infections. Clinically, hypothermia is an indicator for severe infections analogous to hyperthermia, or fever. In fact, neonatal hypothermia is associated in an unclear direction of causality with various pathologies such as surfactant inactivation, increased morbidity from infection, abnormal coagulation, delayed readjustment from the fetal to newborn circulation, hyaline membrane disease and intraventricular hemorrhage in LBW infants.[23]

In contrast, mild therapeutic hypothermia has emerged as a neuroprotective strategy in the treatment of hypoxic ischemic encephalopathy. Recent randomized controlled trials have shown that therapeutic hypothermia initiated within 6 h of birth reduces death and disability in these infants.[24] Induced under controlled clinical conditions, therapeutic hypothermia has been discussed as

being beneficial and outweighing the adverse effects in term newborns with hypoxic ischemic encephalopathy[25] and during or after cardiac surgery.[26]

Neonatal hyperthermia

Heat is transferred *in utero* via the placenta through umbilical arterial blood flow and via the uterus through amniotic fluid to the fetus.[14] At birth, fetal temperature is usually 0.5 to 1.0 °C higher than the mother's[27] and increases not only with elevated maternal temperatures due to prolonged labor, prolonged rupture of the membranes or other infectious etiologies (chorioamnionitis, urinary tract infection, and so on), but also with nulliparity and epidural analgesia.[28]

The most common cause of elevation of body temperature in the newborn is dehydration.[29] Rehydration is both therapeutic and diagnostic if the newborn improves. Elevated temperatures in the neonate rarely reflect intrauterine or perinatal infections. Among the 1 to 2.5% of newborns presenting with hyperthermia, <10% have culture-proven sepsis.[10] In septic newborns, temperature instability more frequently presents as hypothermia. The exact mechanisms that lead to fever in some septic neonates and normal body temperatures in others are ill understood. Infection is thought to produce fever mediated through cytokines such as interleukin1. Antipyretics are effective in reducing the temperature by modifying the central set-point of the hypothalamus. In hyperthermia due to environmental overheating, antipyretics are ineffective, and newborns are appropriately managed by reducing the environmental heat exposure.

Central malformations and intracranial hemorrhages, or congenital pathologies such as the Crisponi syndrome,[30] are rare causes of newborn hyperthermia.

Where thermoprotective devices are used, inappropriate incubation and exposure to radiant warmers are common causes of neonatal hyperthermia,[10] especially when makeshift apparatuses such as light bulbs, hot stones, and so on, are used. These are usually not designed and tested for safety and efficiency, and we discourage their use in favor of skin-to-skin care (SSC).

Management of newborn hypothermia

Qualitative inquiries into current thermal practices

Although lack of equipment is a problem for high-risk neonates in resource-poor settings, knowledge of hypothermia diagnosis and management is another concern. For example, only about half of 160 surveyed health care professionals in India could define neonatal hypothermia correctly or considered it a significant problem, and <20% knew how to correctly record a newborn's temperature.[31] A multinational study showed that knowledge on thermal control, especially concerning the physiology of thermoregulation and criteria for defining hypothermia, was insufficient and thermal control practices were frequently inadequate.[23]

Qualitative research on newborn care can help shed light on the beliefs and attitudes underlying potentially detrimental or harmful practices. Most published studies indicate that high-risk home delivery and newborn care practices that lead to heat loss, such as insufficient heating of the birthplace, placing of the uncovered newborn on the ground or other cold surfaces, delayed wrapping—partly with unclean clothes—and early bathing, remain common in resource-limited settings both in rural and urban areas, in facilities and during home births.[5,32,33]

Heating the birthplace is a critical issue for home births. Studies from Nepal reported that the birthplace was heated in only slightly over half of the settings,[34] often only after birth.[35] Wrapping the child prevents heat loss from evaporation, whereas bathing promotes heat loss. Less than half (46%) of the babies were wrapped within the first 10 min after birth, and almost all of them were bathed within 10 min (89%) or half an hour (96%) after birth.[34] In another study, only 64% of the babies were observed to be wrapped within half an hour after birth, and almost all were bathed within 6 h after birth.[35]

In a study from Tanzania, the practice of bathing newborns immediately after delivery was shown to be motivated by concerns about 'ritual pollution'.[36] In Ghana, early bathing was linked to reducing body odor in later life, shaping the baby's head, and helping the baby sleep and feel clean, and informants felt that changing bathing behaviors would be difficult, especially as babies are bathed early in facilities.[37] A study from Dhaka, Bangladesh, explained that babies are typically bathed soon after birth to purify them from the birth process.[38] Several studies, from Uganda,[39] Ghana[37] and India[40] suggested that in the absence of health facilities prepared to deliver essential newborn care, community members would accept thermoprotective practices such as SSC.

Clinical presentation

It has been estimated that prompt recognition of hypothermia and re-warming of hypothermic infants will avert up to 40% of neonatal deaths.[41] Newborn hypothermia presents with a combination of low core temperature and cold skin, pallor (acrocyanosis), tachypnea (respiratory distress), hypotonia, lethargy or irritability, poor feeding or vomiting. The non-specific clinical presentation and the complex process of thermoregulation discussed above imply a number of differential diagnoses such as infectious etiologies, respiratory distress syndrome, intraventricular hemorrhage or other central nervous causes, hypoglycemia, endocrine causes, or (maternal) drug side effects. Other factors potentially underlying hypothermia include prematurity, cardiovascular diseases and other congenital anomalies.

Diagnosis

Initial assessment should include a history of the baby's exposure to cold and whether the baby has been appropriately clothed and protected.[18] Although a rectal digital thermometer is used in many studies as standard method to measure a newborn's core temperature, this measurement site is associated not only with discomfort and disturbance to the newborn, but also with risks such as rectal perforation and vagal stimulation with resulting arrhythmias, bradycardia and apnea.[42]

The axilla is a less invasive, alternative site that provides reasonably accurate measurements. Mercury-in-glass, gallium-in-glass, digital thermometer, analogous electric thermometer, chemical thermometer and infrared thermometer are all accurate instrument options, with the latter being less hazardous and quicker than the former.[43] Most developed institutions use tympanic thermometers, which have recently shown to be a quick and accurate method to measure a newborn's body temperature,[44] whereas simple rectal thermometers are used in most resource-limited settings. WHO recommends frequent measurements, from every hour in a seriously ill baby, two to four times per day in a small or very small baby, to once daily in an infant progressing well.[18] However, due to their cost, thermometers are often not available in low-resource environments. Moreover, illiteracy and inability to read Arabic numbers have been a challenge to thermometer use.

In the absence of a measurement device, human touch of feet and abdomen has been used as a proxy for body temperature. Studies in India and Nepal have shown human touch to be reasonably reliable for the detection of hypothermia when health workers were trained for these investigations.[45–49] Mothers, however, seem to have a far lower sensitivity than health workers. Only 24% of mothers in India were able to correctly identify hypothermia.[50]

A device based on color indicators developed to detect hypothermia without the use of thermometers was previously found to accurately indicate hypothermia when used by health workers or mothers.[51,52] Its usefulness for some parts of the developing world and the feasibility for illiterate health workers to read the device have been debated.[53,54]

Therapeutic goals of thermal care

The therapeutic goal of thermal care is to keep the newborn in the thermoneutral zone, or thermal neutrality, the environmental temperature range in which the organism has least oxygen consumption.[9] No single environmental temperature is optimal for all babies. In general, the smaller and more premature a newborn is, the less its ability to regulate cold and heat. The optimal environmental temperature thus depends on the maturity (usually estimated by the gestational age) and age of the newborn. Weight, body temperature and skin perfusion as well as clothing of the infant and air humidity also factor in, so that the optimal environmental temperature can be hard to determine. It is narrow, especially in LBW or sick babies, and generally ranges from 32 to 36 °C.[55] It follows that a temperature appropriate for a healthy term baby can be too cold for a preterm infant (and, conversely, what is appropriate for the preterm infant can be too warm for the term baby). In general, most newborns at birth, if left wet and naked, cannot tolerate an environmental temperature of <32 °C. However, if the baby is immediately dried, put skin-to-skin with the mother and covered, the delivery room temperature can be as low as 25 to 28 °C.[17]

Prevention of newborn hypothermia

WHO recognizes maintaining a normal body temperature as a primary principle of newborn care and recommends thermal protection for all infants, with special attention for sick, premature, or small for gestational age infants, for example, <2.5 kg at birth or born before 37 weeks gestation.[18] Several methods can be used for warming the baby and maintaining the baby's body temperature (Table 1). The WHO proposes a 'warm chain', a set of 10 interlinked procedures carried out at birth and during the following hours and days. To be implemented in institutions and (in an abridged form) at home, the warm chain aims to minimize the risk of hypothermia in newborns, and includes warming the delivery room, immediate drying, SSC, early and exclusive breast-feeding, postponing bathing, appropriate clothing and bedding, placing mother and baby together, and in institutions warm transportation, warm resuscitation, and training and awareness raising.[17]

WHO recommends warming the delivery place in preparation for a birth (to at least 25 °C) and to keep the birthplace free from draughts. After delivery, it is crucial to devote some attention to the baby. The first and most substantial heat loss occurs through evaporation of amniotic fluid. Therefore, at birth, recommended first steps to prevent hypothermia are to immediately dry and cover the newborn, even before the cord is cut. While being dried, the baby should be on a warm surface, preferably the mother's chest or abdomen in skin-to-skin contact. The infant should then be clothed or covered, especially the head,[56] and kept in a warm environment, again usually best with the mother. Bathing should be delayed. Draughts, cold surfaces or nearby cold sources such as windows or walls should be avoided as they contribute to heat loss via convection and radiation.

Early breastfeeding, ideally within an hour after delivery, should be encouraged if possible and if not contraindicated. SSC with the mother, for LBW infants also known as kangaroo mother care, most of the time is appropriate to ensure thermal protection of the baby.[57] It requires minimal instructions and, when culturally accepted, can relatively easily be applied even in a community or home setting.[58]

Table 1 Management of hypothermia, modified from WHO 2003[18]

Method	*Selection and use*	*Advantages*	*Risks or disadvantages*	*Availability*
SSC by mother or other person	Management of stable babies with moderate hypothermia, and for hypothermia prevention	Mother can closely monitor other person can provide SSC	Not appropriate for critical conditions	Home and institutional
Kangaroo mother care with LBW babies	Management of stable LBW babies weighing 1.5–2.5 kg, and for hypothermia prevention	Mother can closely monitor. Usually effective to maintain normal body temperature	Mother might not be available. Not for very LBW	Home and institutional
Radiant warmer, water mattress	Management of sick babies weighing $\geqslant$2.5 kg. For initial assessment, treatment and procedures; for hypothermia prevention	Allows observation of baby. Allows for procedures to be performed	Hyperthermia, dehydration. Expensive to procure, requires electricity	Institutional
Incubator	Management of sick or at-risk babies. Continuous care for babies weighing $\leqslant$1.5 kg	Maintains constant temperature and humidity. Easy provision of oxygen. Allows observation of baby (can be naked)	Hyperthermia, dehydration; microbiological contamination. Expensive to procure, maintain, clean; requires electricity. Separation of mother and child	Institutional
Warm room	Recovering babies		Hypothermia. Uncomfortable for adults	Home and institutional
Other methods (water bottles, bricks, and so on)	For emergency situations only	Not recommended	Hypothermia. Hyperthermia, burns	Home and institutional Not recommended

Abbreviations: LBW, low birth weight; SSC, skin-to-skin care; WHO, World Health Organization.

Treatment of newborn hypothermia

According to current WHO guidelines for the treatment of cold babies, moderate hypothermia should be treated by SSC.[18] In severe hypothermia, rewarming the baby with an appropriate and available method in a health care facility setting is warranted, as close monitoring of vital signs including temperature and respiratory rate are essential parts of the management. Blood glucose should be controlled and hypoglycemia under 45 mg dl^{-1} (2.6 mmol l^{-1}) should be treated accordingly. When treating for sepsis, all IV fluids should be given warm. The infant can be discharged once a stable normal temperature is sustained and there are no other issues. Upon discharge, the mother should be counseled to prevent hypothermia at home as discussed above.

There is a relative scarcity of data documenting the effects of recommended thermal newborn care. A recent meta-analysis showed that SSC in conjunction with breastfeeding and recognition of danger signs substantially reduced neonatal mortality in hospital-born preterm babies (birth weight $<$2000 g) in hospital, and was highly effective in reducing severe morbidity, particularly from infection.[59] A study from Western India, in which 36.9% of hospitalized newborns were hypothermic, reported a decrease in this rate to 3.9% with kangaroo mother care.[60] However, in many countries there is resistance from health professionals, mothers and families related to local cultural practices.[61] Although evidence on the effectiveness of SSC in community-based settings is scarce,[40,62] it is estimated that SSC can avert up to 20% of newborn deaths.[63]

In the large Gadchiroli trial in India on home-based neonatal care assessing the outcome of sepsis management, case management included thermal protection of the newborn, and health care workers were given a thermometer, baby clothes and head cover, a blanket and a sleeping bag. Although the study included other interventions and was not specifically designed to prove a particular effect for hypothermia management, it showed a reduction in neonatal and infant mortality by nearly 50% among a malnourished, illiterate and rural study population.[64] A study from Nepal that found a high incidence of hypothermia suggests that simple interventions including immediate drying and another treatment (breast contact, radiant heater and mustard oil massage, or swaddling with an inner layer of plastic wrap) could lower the incidence of hypothermia 2 h after birth from 78 to 23% and 24 h after birth from 49 to 18%.[65] In Zambia, we recently showed that training traditional birth attendants in newborn care with special emphasis on resuscitation and simple thermal protection (wiping the newborn dry and wrapping the dried infant in a separate piece of cloth) along with an intervention to provide early treatment of possible sepsis reduced mortality rates at day 28 after birth by 45%.[66]

Low-cost, low-tech treatment of newborn hypothermia

The use of incubators for thermal protection of newborns has been reported for more than 150 years, since the Parisian obstetrician Jean Louis Paul Denucé in 1857 engineered his couveuse, a device

for the care of a premature infant. In 1878, his local colleague Stéphane Tarnier, using a modified warming chamber for the rearing of poultry, found a decrease in neonatal death rate from 66 to 38% among infants with birth weights <2000 g.[67]

Today, postnatal care devices (isolettes or infant warmers) combine the features of incubators and radiant warmer beds and have evolved with many features, including automated temperature and humidity regulations,[68] oxygen supplementation and light therapy.[11,69] Although beneficial in resource-replete settings,[70] utilization of their complex features requires electricity, concentrated oxygen supply, centralized suction and ongoing skilled maintenance. Priced at about US$15 000 to $36 000,[71] these devices are not affordable for most of the developing world. Simplified versions, such as water-filled mattresses or Indian made, low-cost radiant warmers are power-dependent and not appropriate for resource-limited settings. Polyethylene occlusive skin wrapping is a useful and effective method for delivery room management,[72] but mostly limited to immediate post-delivery care and protection during transport.

A number of postnatal care devices for resource-limited settings are currently in development, some including more sophisticated temperature and humidity regulations. Examples are the 'mkat,' 'Life Raft Incubator,' and 'Neo.nurture', projected to be priced between US $200 and US $625 per unit.[73] The 'Embrace Global', projected to be priced at a US $25 price, is a life vest style incubator in a 'sleeping bag' design.[74] The heat source is a pouch-containing phase change material, which keeps its temperature relatively constant over an extended period of time. The pouch is warmed electrically or by the user simply pouring hot water into a compartment, upon indication by a thermal strip. It can fully open to double as a heat mattress. With some models electricity independent, it can be used both at the institutional and community levels, and serve as visual reminders to mothers and other caretakers, birth attendants and health care workers. Devices such as these, although more costly than education alone, might thus foster improved hypothermia management by transporting a behavioral message to the end user, for example, promoting SSC. Distributed commercially or donated, they could help to raise awareness and enhance perception of the burden of newborn hypothermia.

Conclusions

Thermal protection of the newborn can relatively easily be achieved by warming of the delivery room, immediate drying, wrapping the infant after birth and keeping him or her in close contact with the mother, that is, kangaroo mother care or SSC, immediate and frequent exclusive breastfeeding, delaying of bathing until the infant is physiologically stable, and appropriate warm clothing. These behavior steps represent simple, low-cost measures, which should be integrated into holistic mother and child health packages.

Birth practices even in high-risk environments remain poor, so that interventions must primarily focus on participatory education about hygiene, infection prevention and management, as well as practices to avoid hypothermia. Low-cost, low-technology devices might be helpful in supporting and implementing these practices. Clinical effectiveness and implementation trials will have to investigate which intervention packages and messages for the thermal protection for newborns work best in a given environment, and how to optimally integrate them into existing maternal and newborn health programs.

Conflict of interest

The authors declare no conflict of interest.

References

1 Lyon AJ. Applied physiology: temperature control in the newborn infant. *Curr Pediatr* 2004; **14**: 7.
2 Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG *et al.* Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 2010; **375**: 1969–1987.
3 Lawn JE, Cousens S, Zupan J. 4 Million neonatal deaths: when? where? why? *Lancet* 2005; **365**: 891–900.
4 Kumar V, Shearer JC, Kumar A, Darmstadt GL. Neonatal hypothermia in low resource settings: a review. *J Perinatol* 2009; **29**: 401–412.
5 Mullany LC. Neonatal hypothermia in low-resource settings. *Semin Perinatol* 2010; **34**: 426–433.
6 Takayama JI, Teng W, Uyemoto J, Newman TB, Pantell RH. Body temperature of newborns: what is normal? *Clin Pediatr (Phila)* 2000; **39**: 503–510.
7 Silverman WA, Agate Jr FJ, Fertig JW. A sequential trial of the nonthermal effect of atmospheric humidity on survival of newborn infants of low birth weight. *Pediatrics* 1963; **31**: 719–724.
8 Silverman WA, Fertig JW, Berger AP. The influence of the thermal environment upon the survival of newly born premature infants. *Pediatrics* 1958; **22**: 876–886.
9 Hey E. Thermal neutrality. *Br Med Bull* 1975; **31**: 69–74.
10 Baumgart S. Iatrogenic hyperthermia and hypothermia in the neonate. *Clin Perinatol* 2008; **35**: 183–197, ix-x.
11 Sherman TI, Greenspan JS, St Clair N, Touch SM, Shaffer TH. Optimizing the neonatal thermal environment. *Neonatal Netw* 2006; **25**: 251–260.
12 Charpak N, Ruiz-Pelaez JG, Charpak Y. Rey-martinez Kangaroo mother program: an alternative way of caring for low birth weight infants? one year mortality in a two cohort study. *Pediatrics* 1994; **94**: 804–810.
13 Knobel R, Holditch-Davis D. Thermoregulation and heat loss prevention after birth and during neonatal intensive-care unit stabilization of extremely low-birthweight infants. *J Obstet Gynecol Neonatal Nurs* 2007; **36**: 280–287.
14 Asakura H. Fetal and neonatal thermoregulation. *J Nippon Med Sch* 2004; **71**: 360–370.
15 Hackman PS. Recognizing and understanding the cold-stressed term infant. *Neonatal Netw* 2001; **20**: 35–41.
16 Tourneux P, Libert JP, Ghyselen L, Leke A, Delanaud S, Degrugilliers L *et al.* Echanges thermiques et thermoregulation chez le nouveau-ne [Heat exchanges and thermoregulation in the neonate]. *Arch Pediatr* 2009; **16**: 1057–1062.

17 WHO. Thermal protection of the newborn: a practical guide. *Maternal and Newborn Health/Safe Motherhood Unit, Division of Reproductive Health*. World Health Organization: Geneva, 1997.
18 WHO. Managing newborn problems: a guide for doctors, nurses, and midwives. *Integrated Management of Pregnancy and Childbirth*. World Health Organization: Geneva, 2003.
19 Aylott M. The neonatal energy triangle. Part2: Thermoregulatory and respiratory adaptation. *Paediatr Nurs* 2006; **18**: 38–42.
20 Pal DK, Manandhar DS, Rajbhandari S, Land JM, Patel N, de LCAM. Neonatal hypoglycaemia in Nepal 1. Prevalence and risk factors. *Arch Dis Child Fetal Neonatal Ed* 2000; **82**: F46–F51.
21 Dagan R, Gorodischer R. Infections in hypothermic infants younger than 3 months old. *Am J Dis Child* 1984; **138**: 483–485.
22 Goldsmith JR, Arbeli Y, Stone D. Preventability of neonatal cold injury and its contribution to neonatal mortality. *Environ Health Perspect* 1991; **94**: 55–59.
23 Dragovich D, Tamburlini G, Alisjahbana A, Kambarami R, Karagulova J, Lincetto O *et al.* Thermal control of the newborn: knowledge and practice of health professional in seven countries. *Acta Paediatr* 1997; **86**: 645–650.
24 Pfister RH, Soll RF. Hypothermia for the treatment of infants with hypoxic-ischemic encephalopathy. *J Perinatol* 2010; **30**(Suppl): S82–S87.
25 Jacobs S, Hunt R, Tarnow-Mordi W, Inder T, Davis P. Cooling for newborns with hypoxic ischaemic encephalopathy. *Cochrane Database Syst Rev* 2007; CD003311.
26 Dominguez TE, Wernovsky G, Gaynor JW. Cause and prevention of central nervous system injury in neonates undergoing cardiac surgery. *Semin Thorac Cardiovasc Surg* 2007; **19**: 269–277.
27 Lyon A. Temperature control in the neonate. *Paediatr Child Health* 2008; **18**: 155–160.
28 Laptook AR, Watkinson M. Temperature management in the delivery room. *Semin Fetal Neonatal Med* 2008; **13**: 383–391.
29 Maayan-Metzger A, Mazkereth R, Kuint J. Fever in healthy asymptomatic newborns during the first days of life. *Arch Dis Child Fetal Neonatal Ed* 2003; **88**: F312–F314.
30 Herholz J, Crisponi L, Mallick BN, Rutsch F. Successful treatment of cold-induced sweating in Crisponi syndrome and its possible mechanism of action. *Dev Med Child Neurol* 2010; **52**: 494–497.
31 Choudhary SP, Bajaj RK, Gupta RK. Knowledge, attitude and practices about neonatal hypothermia among medical and paramedical staff. *Indian J Pediatr* 2000; **67**: 491–496.
32 Ayaz A, Saleem S. Neonatal mortality and prevalence of practices for newborn care in a squatter settlement of Karachi, Pakistan: a cross-sectional study. *PLoS One* 2010; **5**: e13783.
33 Ayaz S, Efe SY. Potentially harmful traditional practices during pregnancy and postpartum. *Eur J Contracept Reprod Health Care* 2008; **13**: 282–288.
34 Sreeramareddy CT, Joshi HS, Sreekumaran BV, Giri S, Chuni N. Home delivery and newborn care practices among urban women in western Nepal: a questionnaire survey. *BMC Pregnancy Childbirth* 2006; **6**: 27.
35 Osrin D, Tumbahangphe KM, Shrestha D, Mesko N, Shrestha BP, Manandhar MK *et al.* Cross sectional, community based study of care of newborn infants in Nepal. *BMJ* 2002; **325**: 1063.
36 Thairu L, Pelto G. Newborn care practices in Pemba Island (Tanzania) and their implications for newborn health and survival. *Matern Child Nutr* 2008; **4**: 194–208.
37 Hill Z, Tawiah-Agyemang C, Manu A, Okyere E, Kirkwood BR. Keeping newborns warm: beliefs, practices and potential for behaviour change in rural Ghana. *Trop Med Int Health* 2010; **15**: 1118–1124.
38 Moran AC, Choudhury N, Uz Zaman Khan N, Ahsan Karar Z, Wahed T, Faiz Rashid S *et al.* Newborn care practices among slum dwellers in Dhaka, Bangladesh: a quantitative and qualitative exploratory study. *BMC Pregnancy Childbirth* 2009; **9**: 54.
39 Waiswa P, Nyanzi S, Namusoko-Kalungi S, Peterson S, Tomson G, Pariyo GW. I never thought that this baby would survive; I thought that it would die any time: perceptions and care for preterm babies in eastern Uganda. *Trop Med Int Health* 2010; **15**: 1140–1147.
40 Darmstadt GL, Kumar V, Yadav R, Singh V, Singh P, Mohanty S *et al.* Introduction of community-based skin-to-skin care in rural Uttar Pradesh, India. *J Perinatol* 2006; **26**: 597–604.
41 Darmstadt GL, Walker N, Lawn JE, Bhutta ZA, Haws RA, Cousens S. Saving newborn lives in Asia and Africa: cost and impact of phased scale-up of interventions within the continuum of care. *Health Policy Plan* 2008; **23**: 101–117.
42 Ellis J. Neonatal Hypothermia. *J Neonatal Nurs* 2005; **11**: 7.
43 Cinar ND, Filiz TM. Neonatal thermoregulation. *J Neonatal Nurs* 2006; **12**: 6.
44 Uslu S, Ozdemir H, Bulbul A, Comert S, Bolat F, Can E *et al.* A comparison of different methods of temperature measurements in sick newborns. *J Trop Pediatr* 2011; **57**: 418–423.
45 Singh M, Rao G, Malhotra AK, Deorari AK. Assessment of newborn baby's temperature by human touch: a potentially useful primary care strategy. *Indian Pediatr* 1992; **29**: 449–452.
46 Agarwal S, Sethi V, Pandey RM, Kondal D. Human touch vs axillary digital thermometry for detection of neonatal hypothermia at community level. *J Trop Pediatr* 2007; **54**: 200–201.
47 Ellis M, Manandhar D, Hunt L, Barnett S, Azad K. Touch detection of neonatal hypothermia in Nepal. *Arch Dis Child Fetal Neonatal Ed* 2006; **91**: F367–F368.
48 Tuitui RL, Suwal SN, Shrestha S. Hand-touch method for detection of neonatal hypothermia in Nepal. *J Trop Pediatr* 2010; **57**: 236–238.
49 Agarwal S, Sethi V, Srivastava K, Jha P, Baqui AH. Human touch to detect hypothermia in neonates in Indian slum dwellings. *Indian J Pediatr* 2010; **77**: 759–762.
50 Kumar R, Aggarwal AK. Accuracy of maternal perception of neonatal temperature. *Indian Pediatr* 1996; **33**: 583–585.
51 Morley D, Blumenthal I. A neonatal hypothermia indicator. *Lancet* 2000; **355**: 659–660.
52 Green DA, Kumar A, Khanna R. Neonatal hypothermia detection by ThermoSpot in Indian urban slum dwellings. *Arch Dis Child Fetal Neonatal Ed* 2006; **91**: F96–F98.
53 Kambarami R, Chidede O, Pereira N. ThermoSpot in the detection of neonatal hypothermia. *Ann Trop Paediatr* 2002; **22**: 219–223.
54 Zeal J. ThermoSpot in the detection of neonatal hypothermia. *Ann Trop Paediatr* 2003; **23**: 87–88; author reply.
55 Obladen M. Neugeborenenintensivpflege. *Springer*. Berlin: New York, 2002, 5th edn.
56 Stothers JK. Head insulation and heat loss in the newborn. *Arch Dis Child* 1981; **56**: 530–534.
57 Galligan M. Proposed guidelines for skin-to-skin treatment of neonatal hypothermia. *MCN Am J Matern Child Nurs* 2006; **31**: 298–304; quiz 305–296.
58 Cattaneo A, Davanzo R, Worku B, Surjono A, Echeverria M, Bedri A *et al.* Kangaroo mother care for low birthweight infants: a randomized controlled trial in different settings. *Acta Paediatr* 1998; **87**: 976–985.
59 Lawn JE, Mwansa-Kambafwile J, Horta BL, Barros FC, Cousens S. Kangaroo mother care' to prevent neonatal deaths due to preterm birth complications. *Int J Epidemiol* 2010; **39**(Suppl 1): i144–i154.
60 Suman RP, Udani R, Nanavati R. Kangaroo mother care for low birth weight infants: a randomized controlled trial. *Indian Pediatr* 2008; **45**: 17–23.
61 Charpak N, Ruiz-Pelaez JG. Resistance to implementing kangaroo mother care in developing countries, and proposed solutions. *Acta Paediatr* 2006; **95**: 529–534.
62 Sloan NL, Ahmed S, Mitra SN, Choudhury N, Chowdhury M, Rob U *et al.* Community-based kangaroo mother care to prevent neonatal and infant mortality: a randomized, controlled cluster trial. *Pediatrics* 2008; **121**: e1047–e1059.
63 Acuin CS, Khor GL, Liabsuetrakul T, Achadi EL, Htay TT, Firestone R *et al.* Maternal, neonatal, and child health in southeast Asia: towards greater regional collaboration. *Lancet* 2011; **377**: 516–525.
64 Bang AT, Bang RA, Baitule SB, Reddy MH, Deshmukh MD. Effect of home-based neonatal care and management of sepsis on neonatal mortality: field trial in rural India. *Lancet* 1999; **354**: 1955–1961.
65 Johanson R, Mallab D, Rolfe P, Spencer A. The effect of postdelivery care on neonatal body temperature. *Early Hum Dev* 1990; **21**: 132–133.

66 Gill CJ, Phiri-Mazala G, Guerina NG, Kasimba J, Mulenga C, MacLeod WB *et al.* Effect of training traditional birth attendants on neonatal mortality (Lufwanyama Neonatal Survival Project): randomised controlled study. *BMJ* 2011; **342**: d346.

67 Lussky RC. A century of neonatal medicine. *Minn Med* 1999; **82**: 48–54.

68 Lyon AJ, Oxley C. HeatBalance, a computer program to determine optimum incubator air temperature and humidity. A comparison against nurse settings for infants less than 29 weeks gestation. *Early Hum Dev* 2001; **62**: 33–41.

69 Lyon AJ, Freer Y. Goals and options in keeping preterm babies warm. *Arch Dis Child Fetal Neonatal Ed* 2011; **96**: F71–F74.

70 Soll RF. Heat loss prevention in neonates. *J Perinatol* 2008; **28**(Suppl 1): S57–S59.

71 Maternal and neonatal directed assessment of technology. Infant warmers to prevent neonatal hyothermia. 2011 (accessed on 28 September 2011 at http://mnhtech.org/technology/technology-briefs/infant-warmers/).

72 Knobel RB, Wimmer Jr JE, Holbert D. Heat loss prevention for preterm infants in the delivery room. *J Perinatol* 2005; **25**: 304–308.

73 Design that Matters. Incubators. 2011 (accessed on 28 September 2011 at http://www.designthatmatters.org/news/dtm-blog/project/incubator/).

74 Embrace. embrace global—an infant thermoregulator designed for extreme affordability. 2011 (accessed on 28 September 2011 at http://embraceglobal.org).

Age of natural menopause and atrial fibrillation: The Framingham Heart Study

Jared W. Magnani, MD,[a,b] Carlee B. Moser, MA,[c] Joanne M. Murabito, MD, ScM,[a,d] Kerrie P. Nelson, PhD,[c] Joao D. Fontes, MD,[a] Steven A. Lubitz, MD, MPH,[e,f] Lisa M. Sullivan, PhD,[c] Patrick T. Ellinor, MD, PhD,[e,f] and Emelia J. Benjamin, MD, ScM[a,b,g,h] *Framingham, Boston, and Charlestown, MA*

Background Early menopausal age is associated with risk of cardiovascular events including myocardial infraction, stroke, and increased mortality. Relations between menopausal age and atrial fibrillation (AF) have not been investigated. We examined the association between menopausal age and AF.

Methods Framingham Heart Study women ≥60 years old without prevalent AF and natural menopause were followed up for 10 years or until incident AF. Menopausal age was modeled as a continuous variable and by categories (<45, 45-53, and >53 years). We used Cox proportional hazards regression to determine associations between menopausal age and AF risk.

Results In 1,809 Framingham women (2,662 person-examinations, mean baseline age 71.4 ± 7.6 years, menopausal age 49.8 ± 3.6 years), there were 273 unique participants with incident AF. We did not identify a significant association between the SD of menopausal age (3.6 years) and AF (hazard ratio [HR] per SD 0.94, 95% CI 0.83-1.06; $P = .29$). In a multivariable model with established risk factors for AF, menopausal age was not associated with incident AF (HR per SD 0.97, 95% CI 0.86-1.09; $P = .60$). Examining categorical menopausal age, earlier menopausal age (<45 years) was not significantly associated with increased AF risk compared with older menopausal age >53 years (HR 1.20, 95% CI 0.74-1.94; $P = .52$) or menopausal age 45 to 53 years (HR 1.38, 95% CI 0.93-2.04; $P = .11$).

Conclusion In our moderate-sized, community-based sample, we did not identify menopausal age as significantly increasing AF risk. However, future larger studies will need to examine whether there is a small effect of menopausal age on AF risk. (Am Heart J 2012;163:729-34.)

Younger age at menopause has been associated with increased risk of myocardial infarction, stroke, and mortality.[1,2] In the Nurse's Health Study, menopausal age <40 years and early menopause (age 40-44 years) were associated with up to 1.5- and 1.4-fold increased risk of cardiovascular events compared with menopause ≥55 years.[3] Conversely, older menopausal age (≥53 years) has been related to decreased mortality secondary to ischemic heart disease.[4] Prospective cohort studies have further related earlier age of menopause to increased risk for all-cause mortality.[2,5,6]

Given the association between menopausal age and cardiovascular events, we sought to examine the relation between age of menopause and atrial fibrillation (AF). Atrial fibrillation has profound social and medical burdens, increasing mortality and eliminating the survival advantage that women have over men.[7] Identifying risk factors for AF in women therefore has significant public health importance.[8] To our knowledge, the association between AF and age of menopause has had limited investigation. We considered that the myriad endocrinologic and vascular changes accompanying menopause would predispose women toward increased AF risk. We consequently hypothesized an increased risk of AF for women experiencing menopause at a younger age and, in particular, that cardiac events may mediate the increased risk for developing AF.

From the [a]National Heart, Lung and Blood Institute's and Boston University's Framingham Heart Study, Framingham, MA, [b]Section of Cardiovascular Medicine, Department of Medicine, Boston University, Boston, MA, [c]Department of Biostatistics, Boston University School of Public Health, Boston, MA, [d]Section of General Internal Medicine, Department of Medicine, Boston University, Boston, MA, [e]Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, MA, [f]Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, [g]Preventive Medicine Section, Department of Medicine, Boston University School of Medicine, Boston, MA, and [h]Epidemiology Department, Boston University School of Public Health, Boston, MA.

Dr. Magnani is supported by American Heart Association Award 09FTF219028. This work was supported by grants from the NIH to Drs. Benjamin and Ellinor (HL092577), Dr. Benjamin (RO1AG028321, RC1-HL01056, 1R01HL102214) and Dr. Ellinor (5R21DA027021, 1RO1HL104156, 1K24HL105780) and 6R01-NS17950, N01-HC25195. This work was partially supported by the Evans Center for Interdisciplinary Biomedical Research ARC on "Atrial Fibrillation at Boston University (http://www.bumc.bu.edu/evanscenteribr/).

Submitted November 16, 2011; accepted January 9, 2012.

Reprint requests: Jared W. Magnani, MD, Section of Cardiovascular Medicine, Boston University School of Medicine, 88 E. Newton Street, Boston, MA 02118.

E-mail: jmagnani@bu.edu

0002-8703/$ - see front matter

doi:10.1016/j.ahj.2012.01.010

Methods

Study sample

The Framingham Heart Study is a longitudinal, community-based study designed to investigate cardiovascular disease and its risk factors.[9] Original cohort participants were enrolled in the Framingham Heart Study starting in 1948 and attend examinations every 2 years. Offspring cohort participation started in 1971 and consists of examinations every 4 to 8 years. To have 10-year follow-up, the present study used data from women attending the original cohort examination cycles 11 (1968-1971), 17 (1981-1984), and 23 (1992-1996), and offspring cohort examination cycles 1 (1971-1975), 3 (1983-1987), and 6 (1995-1998). Participants were included in the present analysis when they reached 60 years or older. Age 60 years was used as the minimum age because women were expected to have reached menopause by that age. Before exclusions, the sample was composed of 4,159 participant examinations. Participants were excluded if they had prevalent AF (n = 180), that is, been diagnosed with AF before study entry at 60 years of age; unreliable age of natural menopause secondary to oophorectomy with or without hysterectomy (n = 1,190); unknown age of menopause (n = 1); menopausal age <40 years (deemed premature ovarian failure,[10] n = 39); estrogen use before menstrual cessation (n = 17); or lacked complete data on key AF risk factors (n = 70).

We used cross-sectional pooling to construct the data set consistent with previous Framingham Heart Study analyses of AF,[11] such that study participants without incident AF included in the analysis were eligible to reenter the analysis for subsequent 10-year intervals. Study participants were followed up prospectively for incident AF for a maximum duration of 10 years after each baseline examination. Participants provided written informed consent at each examination. Study protocols and all examination cycles were approved by the institutional review board of the Boston University Medical Center.

Clinical assessments

Participants underwent a physician-administered medical interview, history, and examination at each Framingham Heart Study examination.[12] Body mass index was calculated from weight in kilograms divided by height in square meters and systolic blood pressure as the mean of 2 seated measurements obtained during a standardized examination. Hypertension treatment was established by self-report of prescribed medications. Heart murmurs of clinical significance were scored as at least grade 3 of 6 systolic or any diastolic murmur recorded by a Framingham Heart Study physician at the standardized examination. The electrocardiographic (ECG) PR interval was calculated from the 12-lead resting ECG as previously described.[13] Heart failure was adjudicated by 3 Framingham Heart Study physicians according to established major and minor criteria.[14] Covariates were selected for inclusion in the present analysis from a previously published AF risk prediction model.[11] Atrial fibrillation was determined by presence of AF or atrial flutter on ECG or Holter monitoring obtained at a Framingham Heart Study examination, external clinician visit, or during hospitalization, and all available outside visits to clinicians for cardiovascular diagnoses. Incident AF was adjudicated by at least 2 Framingham Heart Study cardiologists.

Age of menopause

Age of natural menopause was established by a standardized, physician-administered interview at each examination. Women were queried about their menstrual status, whether periods had stopped for 1 year or more; age periods stopped, cause, defined as natural, surgical or other; history of gynecologic surgery (hysterectomy and oophorectomy, including number of ovaries removed); and use of hormone-replacement therapies. *Natural menopause* was defined as the natural cessation of menses for ≥1 year. Menopausal age was categorized as age <45, 45 to 53, and >53 years after examining the distribution of menopausal age in a prior Framingham Heart Study analysis.[15] Women with a history of surgical cessation of menstrual periods were excluded because of lack of reliability of ascertaining menopausal age, consistent with prior Framingham Heart Study analyses of menopausal age as an exposure for cardiovascular outcomes.[16]

Statistical analyses

We summarized continuous variables with means, SDs, medians, and interquartile ranges and examined distributions graphically. Distributions of categorical variables were examined by frequency. Menopausal age was modeled as a continuous measure (in years) and using categories defined by age cutpoints: <45, 45 to 53, and >53 years. Atrial fibrillation incidence rates were determined by determining the number of events per menopausal age category per 1000 person-years. The association between the SD of menopausal age and incident AF was determined by Cox proportional hazards regression analysis with censoring at 10 years. Models were adjusted for risk factors associated with AF, including body mass index, systolic blood pressure, hypertension treatment, PR interval, significant murmur, and prevalent heart failure. Age was not included in the model because of its collinearity with age of menopause. Risk factors were measured for participants entering each 10-year risk assessment at the baseline examination. We then constructed cumulative incidence curves using menopausal age as a categorical variable. A 2-sided P value of <.05 was considered statistically significant, and all analyses were conducted using SAS version 9.1 (SAS Institute, Cary, NC).

The authors are solely responsible for the design and conduct of this study, the study analyses, and the drafting and editing of the manuscript, and its final contents.

Results

The study cohort consisted of 1,809 unique Framingham Heart Study participants with 2,662 examinations. The mean age was 71.4 ± 7.6 years, and the mean age at natural menopause was 49.8 ± 3.6 years (Table I). There were 273 incident AF events in follow-up. The unadjusted incidence rates per 1000 person-years were 15.8 in the menopausal age <45 years category, compared with 11.5 and 13.3 in the menopausal age 45 to 53 and >53 years categories (P = .24 between different age categories) (Table II).

When examining menopausal age as a continuous variable, we did not observe a statistically significant

Table I. Characteristics of 2,662* participant examinations included in the analysis

Variable†	Characteristic
Age (y)	71.4 ± 7.6
Age of menopause (y)	49.8 ± 3.6
Body mass index (kg/m^2)	25.9 ± 4.9
Systolic blood pressure (mm Hg)	142 ± 22
Hypertension treatment	447 (17)
PR interval duration (ms)	164 ± 25
Significant murmur‡	142 (5.3)
Heart failure	51 (1.9)

*The 2,662 baseline examinations correspond to pooled examinations from 1,809 unique individuals.
† Data are presented as mean ± SD, or n (%).
‡ Significant murmur defined as grade 3 of 6 systolic or any diastolic murmur.

Table II. Incidence of AF from 2,662* participant examinations in eligible Framingham Heart Study women

	Total events	Incidence rate per 1000 person-years (95% CI)
Menopausal age (y) by category		
Age <45 (n = 235)	28	15.8 (9.9-21.7)
Age 45-53 (n = 2246)	206	11.5 (9.9-13.1)
Age >53 (n = 366)	39	13.3 (9.1-17.5)
Total	273	12.1 (10.7-13.5)

*The 2,662 baseline examinations correspond to pooled examinations from 1,809 unique individuals.

association between menopausal age and risk of incident AF (P = .29). Similarly, when examining menopausal age categories, earlier menopausal age (<45) was not associated with increased AF risk compared with older menopausal age >53 years (P = .52) or when compared with menopausal age 45 to 53 years (P = .11). In multivariable analysis incorporating established risk factors for AF, age of menopause was not significantly associated with incident AF in the above analyses (Table III).

The cumulative incidence curves demonstrated no significant differences in incident AF across the 3 categories of menopausal age over 10-year follow-up (Figure 1). The log-rank test did not reach statistical significance as well (P = .24).

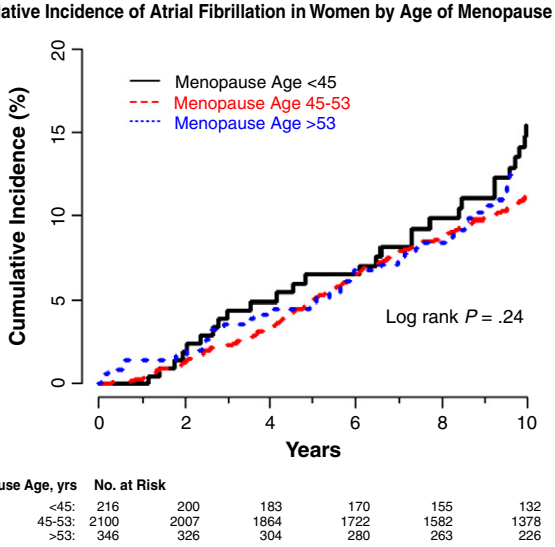
Following the above results, we conducted an analysis to determine the effect size that we were powered to detect. Assuming 80% power and given our number of observed AF events, we would have required a hazard ratio (HR) of 0.84 to observe a significant association between increasing menopausal age and decreased risk of incident AF. In contrast, our observed HR was 0.93.

Table III. Relation of age at menopause to 10-year risk of AF

Variable	HR	95% CI	P
Per SD of menopausal age			
Menopausal age	0.94	0.83-1.06	.29
Multivariable adjusted*	0.97	0.86-1.09	.60
Early menopause (<45) vs late menopause (>53)			
Menopausal age	1.20	0.74-1.94	.52
Multivariable adjusted	1.05	0.64-1.73	.84
Early menopause (<45) vs mid menopause (45-53)			
Menopausal age	1.38	0.93-2.04	.11
Multivariable adjusted	1.15	0.76-1.72	.51

The 2,662 baseline examinations correspond to pooled examinations from 1,809 unique individuals.
*The multivariable covariates are body mass index, hypertension, cardiac murmur, prevalent heart failure, and PR interval.

Figure 1



The figure shows the cumulative incidence of AF, stratified by 3 categories of menopausal age: age <45, age 45 to 53, and age >53 years. The log-rank test shows no significant difference between the 3 curves (P = .24).

Discussion

We hypothesized an association between menopausal age and 10-year risk of incident AF in a prospective, community-based cohort and, specifically, that decreased age of menopause would result in increased incident AF risk. Our hypothesis stemmed from the established association between earlier menopausal age and augmented risk for cardiovascular events.

A second rationale was the cardiovascular adaptation and remodeling by complex vascular and endocrinologic

pathways resulting from menopause.[17] Aging and cardiovascular adaptation after menopause are multifactorial processes that, in turn, yield an increased burden of clinical risk factors associated with AF. Hypertension increases in prevalence as women age,[18] is common across ethnic and racial groups,[19,20] and is a chief risk factor for AF in community-based studies.[21,22] Hypertension has been associated with increased left ventricular hypertrophy and mass in postmenopausal women[23] and increases risk for heart failure.[24] Both hypertension and heart failure are associated with AF.[25] Obesity has risen in prevalence in adults, and in the most recent National Health and Nutrition Examination Survey, 33.6% of women ≥60 years old had a mean body mass index of ≥30 kg/m^2.[26] Menopause results in redistribution of fat depots,[27] including epicardial fat,[28] which has been demonstrated as associated with increased risk for AF.[29] The interaction of such diverse anthropometric and clinical factors with AF risk in postmenopausal women merits continued investigation.

The final rationale for our hypothesis was the association between menopause and risk factors for AF. Inflammatory markers (eg, C-reactive protein) are associated with increased risk of cardiovascular events in women,[30-32] have been related to menopausal status,[33] and similarly have been associated with AF.[34] Postmenopausal hematologic changes include increased hematocrit, greater plasma viscosity, and higher fibrinogen levels, which may enhance risk for cardiovascular disease and AF[35]; in the Women's Health Study, fibrinogen has recently been associated with increased AF risk.[36]

Sex-specific differences in the incidence of AF and cardiac arrhythmias are well described,[37,38] and epidemiologic data have observed that AF is less prevalent in women than men.[39] Incident AF, in general, occurs at an older age in women than men.[21] Sex differences in atrial electrophysiology are demonstrated by decreased P wave indices—a noninvasive assessment of atrial electrophysiologic function—in women compared with men.[40] The study of sex-specific electrophysiologic differences from hormonal influences, specifically estrogen, is ongoing. Estrogen receptors have been studied in animal and human cardiac structures.[41,42] Mice receiving ovariectomies had decreased atrioventricular nodal conduction, as measured by PR and AH intervals, compared with those receiving estrogen or intact animals.[43] In human studies, there is age-associated methylation of the estrogen receptor gene alpha in atrial tissue, suggesting down-regulation of estrogen receptor expression.[44] Modeling of cellular differences in cardiac repolarization has demonstrated female cells having increased potential for arrhythmogenic early after depolarizations compared with male cells.[45]

Further studies are necessary to assess other cardiovascular markers (eg, natriuretic peptides, left atrial parameters, P wave indices), their association with menopausal age, and if they modify AF risk in postmenopausal women. Remaining questions concern estrogen exposure, receptor activity, and modification of atrial electrophysiology across the spectrum of menopause.

Our study has multiple strengths. It was conducted in a community-based cohort with routine examinations occurring every 2 to 8 years. Such frequent contact provided an opportunity for longitudinal assessments and facilitated the cross-sectional pooling used in this study. We included participants only ≥60 years old, thereby verifying that all participants had achieved menopause, and facilitated follow-up in older age when participants were at increased risk for incident AF. Participants with menopausal age <40 years were excluded because of concern for premature ovarian failure. In addition, the Framingham Heart Study's routine collection of varied medical and clinical records yielded the identification of incident AF.

The present study has several limitations. Framingham Heart Study participants were mostly older and primarily of European descent; the generalizability of our findings to younger women and other races and ethnicities is unknown, particularly given racial differences in AF incidence.[46,47] A chief limitation is our lack of power given our sample size and number of incident events. A larger cohort may have a greater number of events, thereby obtaining higher power and the ability to determine smaller risks between menopausal age and incident AF. In fact, we determined that our study would have required approximately 1,400 events in order for the HR of 0.93 to reach statistical significance. Furthermore, we did not examine postmenopausal hormone-replacement therapy and its potential effects on AF risk. Inclusion of hormone use will be essential for larger studies exploring the association between menopause and AF. Our results may further suffer from recall bias in reporting age of natural menopause. However, as we recorded age of menopause before the development of AF, such recall bias is unlikely to be biased by occurrence of AF. Such random misclassification would have biased our results toward the null. It is also possible that women had unrecognized episodes of paroxysmal AF, resulting in misclassification of outcome status. More extensive rhythm monitoring, challenging in a community-based study, would be necessary to capture paroxysmal AF.

To our knowledge, the present study is the first to examine the relation between menopausal age and incident AF. Although we did not observe a significant association between menopausal age and AF risk, studies in larger cohorts may have increased power to explore such a potential association. Atrial fibrillation carries tremendous social and medical burdens, and the number of older adults in the United States continues

to increase. Identification of novel risk factors will serve public health efforts by enhancing risk stratification and prevention initiatives.

References

1. Gordon T, Kannel WB, Hjortland MC, et al. Menopause and coronary heart disease. The Framingham Study. Ann Intern Med 1978;89:157-61.
2. Mondul AM, Rodriguez C, Jacobs EJ, et al. Age at natural menopause and cause-specific mortality. Am J Epidemiol 2005;162: 1089-97.
3. Hu FB, Grodstein F, Hennekens CH, et al. Age at natural menopause and risk of cardiovascular disease. Arch Intern Med 1999;159: 1061-6.
4. Jacobsen BK, Nilssen S, Heuch I, et al. Does age at natural menopause affect mortality from ischemic heart disease? J Clin Epidemiol 1997;50:475-9.
5. Cooper GS, Baird DD, Weinberg CR, et al. Age at menopause and childbearing patterns in relation to mortality. Am J Epidemiol 2000; 151:620-3.
6. Jacobsen BK, Heuch I, Kvale G. Age at natural menopause and all-cause mortality: a 37-year follow-up of 19,731 Norwegian women. Am J Epidemiol 2003;157:923-9.
7. Benjamin EJ, Wolf PA, D'Agostino RB, et al. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. [see comments] Circulation 1998;98:946-52.
8. Benjamin EJ, Chen PS, Bild DE, et al. Prevention of atrial fibrillation: report from a national heart, lung, and blood institute workshop. Circulation 2009;119:606-18.
9. Dawber TR, Meadors GF, Moore Jr FE. Epidemiological approaches to heart disease: the Framingham Study. Am J Public Health Nations Health 1951;41:279-81.
10. Maclaran K, Panay N. Premature ovarian failure. J Fam Plann Reprod Health Care 2011;37:35-42.
11. Schnabel RB, Sullivan LM, Levy D, et al. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. Lancet 2009;373:739-45.
12. Kannel WB, Feinleib M, McNamara PM, et al. An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol 1979;110:281-90.
13. Cheng S, Keyes MJ, Larson MG, et al. Long-term outcomes in individuals with prolonged PR interval or first-degree atrioventricular block. JAMA 2009;301:2571-7.
14. McKee PA, Castelli WP, McNamara PM, et al. The natural history of congestive heart failure: the Framingham study. N Engl J Med 1971; 285:1441-6.
15. Murabito JM, Yang Q, Fox C, et al. Heritability of age at natural menopause in the Framingham Heart Study. J Clin Endocrinol Metab 2005;90:3427-30.
16. Lisabeth LD, Beiser AS, Brown DL, et al. Age at natural menopause and risk of ischemic stroke: the Framingham heart study. Stroke 2009;40:1044-9.
17. Yanes LL, Reckelhoff JF. Postmenopausal hypertension. Am J Hypertens 2011;24:740-9.
18. Hajjar I, Kotchen TA. Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988-2000. JAMA 2003;290:199-206.
19. Wassertheil-Smoller S, Anderson G, Psaty BM, et al. Hypertension and its treatment in postmenopausal women: baseline data from the Women's Health Initiative. Hypertension 2000;36:780-9.
20. Hsia J, Margolis KL, Eaton CB, et al. Prehypertension and cardiovascular disease risk in the Women's Health Initiative. Circulation 2007;115:855-60.
21. Benjamin EJ, Levy D, Vaziri SM, et al. Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study. JAMA 1994;271:840-4.
22. Psaty BM, Manolio TA, Kuller LH, et al. Incidence of and risk factors for atrial fibrillation in older adults. Circulation 1997;96: 2455-61.
23. Oberman A, Prineas RJ, Larson JC, et al. Prevalence and determinants of electrocardiographic left ventricular hypertrophy among a multiethnic population of postmenopausal women (The Women's Health Initiative). Am J Cardiol 2006;97:512-9.
24. Lloyd-Jones DM, Larson MG, Leip EP, et al. Lifetime risk for developing congestive heart failure: the Framingham Heart Study. Circulation 2002;106:3068-72.
25. Wang TJ, Larson MG, Levy D, et al. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. Circulation 2003;107: 2920-5.
26. Flegal KM, Carroll MD, Ogden CL, et al. Prevalence and trends in obesity among US adults, 1999-2008. JAMA 2010;303: 235-41.
27. Ozbey N, Sencer E, Molvalilar S, et al. Body fat distribution and cardiovascular disease risk factors in pre- and postmenopausal obese women with similar BMI. Endocr J 2002;49:503-9.
28. de Vos AM, Prokop M, Roos CJ, et al. Peri-coronary epicardial adipose tissue is related to cardiovascular risk factors and coronary artery calcification in post-menopausal women. Eur Heart J 2008;29: 777-83.
29. Al Chekakie MO, Welles CC, Metoyer R, et al. Pericardial fat is independently associated with human atrial fibrillation. J Am Coll Cardiol 2010;56:784-8.
30. Rifai N, Buring JE, Lee IM, et al. Is C-reactive protein specific for vascular disease in women? Ann Intern Med 2002;136: 529-33.
31. Ridker PM, Hennekens CH, Buring JE, et al. C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. N Engl J Med 2000;342:836-43.
32. Ridker PM, Buring JE, Shih J, et al. Prospective study of C-reactive protein and the risk of future cardiovascular events among apparently healthy women. Circulation 1998;98:731-3.
33. Woodward M, Rumley A, Lowe GD, et al. C-reactive protein: associations with haematological variables, cardiovascular risk factors and prevalent cardiovascular disease. Br J Haematol 2003; 122:135-41.
34. Schnabel RB, Larson MG, Yamamoto JF, et al. Relations of biomarkers of distinct pathophysiological pathways and atrial fibrillation incidence in the community. Circulation 2010;121: 200-7.
35. Woodward M, Rumley A, Tunstall-Pedoe H, et al. Associations of blood rheology and interleukin-6 with cardiovascular risk factors and prevalent cardiovascular disease. Br J Haematol 1999;104: 246-57.
36. Conen D, Ridker PM, Everett BM, et al. A multimarker approach to assess the influence of inflammation on the incidence of atrial fibrillation in women. Eur Heart J 2010;31:1730-6.
37. Schulze-Bahr E, Kirchhof P, Eckardt L, et al. Gender differences in cardiac arrhythmias. Herz 2005;30:390-400.
38. Yarnoz MJ, Curtis AB. More reasons why men and women are not the same (gender differences in electrophysiology and arrhythmias). Am J Cardiol 2008;101:1291-6.

39. Michelena HI, Powell BD, Brady PA, et al. Gender in atrial fibrillation: ten years later. Gend Med 2010;7:206-17.
40. Magnani JW, Johnson VM, Sullivan LM, et al. P-wave indices: derivation of reference values from the Framingham Heart Study. Ann Noninvasive Electrocardiol 2010;15:344-52.
41. Karas RH, Patterson BL, Mendelsohn ME. Human vascular smooth muscle cells contain functional estrogen receptor. Circulation 1994;89:1943-50.
42. Lizotte E, Grandy SA, Tremblay A, et al. Expression, distribution and regulation of sex steroid hormone receptors in mouse heart. Cell Physiol Biochem 2009;23:75-86.
43. Saba S, Zhu W, Aronovitz MJ, et al. Effects of estrogen on cardiac electrophysiology in female mice. J Cardiovasc Electrophysiol 2002; 13:276-80.
44. Post WS, Goldschmidt-Clermont PJ, Wilhide CC, et al. Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system. Cardiovasc Res 1999; 43:985-91.
45. Verkerk AO, Wilders R, de GW, et al. Cellular basis of sex disparities in human cardiac electrophysiology. Acta Physiol (Oxf) 2006;187: 459-77.
46. Alonso A, Agarwal SK, Soliman EZ, et al. Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. Am Heart J 2009;158:111-7.
47. Marcus GM, Olgin JE, Whooley M, et al. Racial differences in atrial fibrillation prevalence and left atrial size. Am J Med 2010;123: 375-7.

Comparing Alcohol Screening Measures Among HIV-Infected and -Uninfected Men

Kathleen A. McGinnis, Amy C. Justice, Kevin L. Kraemer, Richard Saitz, Kendall J. Bryant, and David A. Fiellin

Background: Brief measures of unhealthy alcohol use have not been well validated among people with HIV. We compared the Alcohol Use Disorders Identification Test (AUDIT) to reference standards for unhealthy alcohol use based on 30-day Timeline Follow Back (TLFB) and Composite International Diagnostic Interview—Substance Abuse Module (CIDI-SAM), among 837 male HIV-infected and -uninfected patients in the Veterans Aging Cohort Study.

Methods: Three reference standards were (i) *Risky drinking*—based on TLFB >14 drinks over 7 consecutive days or >4 drinks on 1 day; (ii) *Alcohol dependence*—based on a CIDI-SAM diagnosis; and (iii) *Unhealthy alcohol use*—risky drinking or a CIDI-SAM diagnosis of abuse or dependence. Various cutoffs for the AUDIT, AUDIT-C, and heavy episodic drinking were compared with the reference standards.

Results: Mean age of patients was 52 years, 53% (444) were HIV-infected, and 53% (444) were African American. Among HIV-infected and -uninfected patients, the prevalence of risky drinking (14 vs. 12%, respectively), alcohol dependence (8 vs. 7%), and unhealthy alcohol use (22 vs. 20%) was similar. For risky drinking and alcohol dependence, multiple cutoffs of AUDIT, AUDIT-C, and heavy episodic drinking provided good sensitivity ($\geq$80%) and specificity ($\geq$90%). For unhealthy alcohol use, few cutoffs provided sensitivity $\geq$80%; however, many cutoffs provided good specificity. For all 3 alcohol screening measures, sensitivity improved when heavy episodic drinking was included with the cutoff. Sensitivity of measures for risky drinking and unhealthy alcohol use was lower in HIV-infected than in uninfected patients.

Conclusions: For identifying risky drinking, alcohol dependence, and unhealthy alcohol use, AUDIT-C performs as well as AUDIT and similarly in HIV-infected and -uninfected patients. Cutoffs should be based on the importance of specific operating characteristics for the intended research or clinical use. Incorporating heavy episodic drinking increased sensitivity for detecting alcohol dependence and unhealthy alcohol use.

Key Words: AUDIT, HIV, Veterans, Alcohol-Related Disorders, Mass Screening.

"UNHEALTHY ALCOHOL USE" includes any amount of drinking above "low risk" drinking and specifically includes risky drinking, alcohol abuse, and alcohol dependence (Bradley et al., 2009; Saitz, 2005). Unhealthy alcohol use is common in HIV-infected individuals and contributes to adverse health consequences (Conigliaro et al., 2003; Saitz, 2005). Alcohol consumption can influence HIV disease progression by impacting patients' adherence and response to antiretroviral treatment (Baum et al., 2010; Samet et al., 2006). Despite these findings, and the potential that even low levels of alcohol consumption may cause harm among those infected with HIV (Braithwaite and Bryant, 2010; Sullivan et al., 2011), there are no detailed studies comparing the operating characteristics of standard alcohol screening measures between HIV-infected and -uninfected individuals or validating the use of alcohol screening measures and identifying appropriate cutoffs to use for HIV-infected patients. This creates a problem for researchers and clinicians in deciding which measures should be used to identify and characterize unhealthy alcohol use among HIV-infected individuals.

To conduct meaningful research on alcohol use and related consequences in those with HIV infection and to identify patients who may benefit from counseling or pharmacologic interventions, researchers and clinicians need valid and feasible measures for identifying unhealthy alcohol

From the Center for Health Equity Research and Promotion (KAM), VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania; Division of General Internal Medicine (ACJ, DAF), Yale University School of Medicine, New Haven, Connecticut; Veterans Aging Cohort Study Coordinating Center (ACJ, DAF), West Haven VA Healthcare System, West Haven, Connecticut; Center for Interdisciplinary Research on AIDS (ACJ, DAF), Yale University School of Public Health, New Haven, Connecticut; Department of Medicine (KLK), Center for Research on Health Care, Division of General Internal Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania; Department of Medicine (RS), Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, Massachusetts; Department of Epidemiology (RS), Boston University School of Public Health, Boston, Massachusetts; and National Institute on Alcohol Abuse and Alcoholism (KJB), Bethesda, Maryland.

Received for publication December 7, 2011; accepted July 11, 2012.

Reprint requests: Kathleen McGinnis, MS, Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System, 7180 Highland Drive (151C-H), Pittsburgh, PA 15206-4900; Tel.: 412-215-9927; Fax: 412-365-5264; E-mail: kathleen.mcginnis3@va.gov

The views are not those of the Department of Veterans Affairs or the United States Government.

DOI: 10.1111/j.1530-0277.2012.01937.x

use. For instance, certain pharmacotherapies are indicated primarily for those who meet criteria for alcohol dependence. There are many instruments designed to identify those who may be at increased risk for unhealthy alcohol use, including brief screening measures such as the Alcohol Use Disorders Identification Test (AUDIT; Babor et al., 1989). More time-intensive methods such as the 30-day Timeline Follow Back (TLFB; Sobell et al., 1988a) and the Composite International Diagnostic Interview—Substance Abuse Module (CIDI-SAM; Compton, 1993) provide more definitive information about unhealthy alcohol use. Many prior studies have evaluated the use of the AUDIT for identifying various types of unhealthy alcohol use and have identified appropriate cutoffs to use in various populations (Babor et al., 1989; Bradley et al., 1998a,b; Gordon et al., 2001; Gual et al., 2002; Reinert and Allen, 2002; Rubinsky et al., 2010; Saunders et al., 1993; Schmidt et al., 1995) but not in an HIV-infected population.

The aim of this research was to determine the performance of various screening instruments and thresholds for 3 types of alcohol use in a large cohort of HIV-infected and -uninfected patients receiving primary care. The 3 drinking categories evaluated are (i) risky drinking (Fleming, 2004/2005), (ii) alcohol dependence (Diagnostic and Statistical Manual of Mental Disorders [DSM]), and (iii) unhealthy alcohol use (Saitz, 2005).

MATERIALS AND METHODS

Setting and Subjects

Subjects were male participants of the Veterans Aging Cohort Study (VACS), a prospective multisite observational study focused on the role of alcohol use on health in HIV-infected and -uninfected individuals (Justice et al., 2006a), funded by the National Institute on Alcohol Abuse and Alcoholism. VACS is an ongoing study conducted at 8 Veterans Affairs facilities (Atlanta, Georgia; Baltimore, Maryland; Bronx, New York; Houston, Texas; Los Angeles, California; New York City, New York; Pittsburgh, Pennsylvania; and Washington, District of Columbia). HIV-infected individuals are recruited from Infectious Disease clinics that provide primary medical care at the participating sites. All HIV-infected patients presenting to each participating site are eligible. Uninfected controls are recruited from the General Internal Medicine clinics at the same sites and are matched to the demographics of the HIV-infected subjects on 5-year age blocks, race/ethnicity, and gender. Subjects in this analysis are 837 male VACS participants enrolled from June 2002 to February 2004 who were part of a substudy designed to assess alcohol use and the validity of alcohol screening measures. VACS participants were asked to participate in the substudy if they had reported having at least 1 alcoholic drink in the past year. The study was approved by the Institutional Review Board of each site.

Measures of Alcohol Use Collected in VACS

The AUDIT was collected via self-administered paper survey at the 1-year follow-up (Babor et al., 1989). The CIDI-SAM and 30-day TLFB were administered via telephone interview shortly after the 1-year follow-up paper survey.

The CIDI-SAM operationalizes the diagnostic criteria of DSM-III, DSM-III-R, DSM-IV and International Classification of Diseases—10th Revision (ICD-10) psychoactive substance use disorders (Compton, 1993; Kessler et al., 1998). CIDI-SAM data were used to identify both lifetime and current (past 12 month) alcohol abuse and dependence disorders. The alcohol 30-day TLFB, a validated calendar-based method whereby subjects provide retrospective estimates of their daily drinking was used to collect information on alcohol consumption over the 30 days prior to the telephone interview. The TLFB has been evaluated in clinical and nonclinical populations and is validated when administered in person and by telephone (Maisto et al., 2009; Sobell and Sobell, 1992; Sobell et al., 1988a,b). The 10-item AUDIT (Babor et al., 1989; Gordon et al., 2001; Reinert and Allen, 2002; Saunders et al., 1993; Schmidt et al., 1995) was administered during the 1-year follow-up survey and includes the 3-item AUDIT-C, an instrument that consists of the 3 consumption items of the AUDIT (Bush et al., 1998; Dawson et al., 2005; Gual et al., 2002). The AUDIT-C has similar sensitivities and specificities for identifying unhealthy alcohol use to the AUDIT (Bradley et al., 1998a,b; Gual et al., 2002) yet is shorter.

We used the third item of the AUDIT, which asks: "How often do you have 6 or more drinks on 1 occasion?" with response options "never," "less than monthly," "monthly," "weekly," and "daily or almost daily" to identify patients with heavy episodic drinking. Those who indicated that they consumed 6 or more drinks (chose response option of "less than monthly" or more frequently) were classified as "ever" heavy episodic drinkers. Heavy episodic drinking has been associated with many adverse health effects including unintentional injuries, acute myocardial infarction, ischemic stroke, sexually transmitted diseases (Hansagi et al., 1995; Naimi et al., 2003; Raj et al., 2009), and among those with HIV, it has been associated with decreased medication adherence (Braithwaite et al., 2008). To determine whether we could improve the performance of the AUDIT and AUDIT-C measures with reference standard measures by incorporating the single heavy episodic drinking item from the AUDIT, we also created composite measures that were positive if the AUDIT measure was greater than a specified cutoff *or* if the person reported any heavy episodic drinking in the past year. These measures represent a broader spectrum of unhealthy alcohol use than do an AUDIT cutoff or heavy episodic drinking alone.

We used ICD-9 codes to identify subjects diagnosed with alcohol dependence. We also used ICD-9 codes to identify subjects with an alcohol-related diagnosis, which included ICD-9 codes reflecting alcohol abuse, alcohol dependence, or problems stemming from excessive alcohol consumption (e.g., alcoholic cirrhosis). The alcohol-related diagnosis ICD-9 codes used were based on the work of Piette and are listed on the VACS web site (www.vacohort.org) (Justice et al., 2006b; Piette et al., 1998). Patients were considered to have a diagnosis of alcohol dependence or an alcohol-related diagnosis if they had at least 1 inpatient or 2 outpatient alcohol-related ICD-9 codes ever and within the 12 months prior to and including the date of the follow-up telephone interview. Requiring at least 1 inpatient or 2 outpatient ICD-9 codes has been shown to improve the accuracy of these codes (Justice et al., 2006b).

Alcohol Reference Measures for Risky Drinking, Dependence, and Unhealthy Use

We used established reference standards ("gold standards") for risky drinking, alcohol dependence, and unhealthy alcohol use. The reference standard for *risky drinking* was based on the 30-day TLFB assessment of amounts that increase the risk for health consequences. We defined risky drinking as consuming >14 drinks over any consecutive 7-day period or >4 drinks in 1 day based on the guidelines recommended by the National Institute on Alcohol Abuse and Alcoholism to identify those at risk for alcohol-related problems (Fleming, 2004/2005). The reference standard for past year alcohol *dependence* was based on the CIDI-SAM diagnosis of dependence. We chose to use a dependence-only reference measure instead of dependence/abuse

because certain treatments and pharmacotherapies are indicated primarily for those who meet criteria for alcohol dependence. *Unhealthy alcohol use in the past year* was defined as having a CIDI-SAM diagnosis of alcohol abuse or dependence in the past year or meeting criteria for risky drinking based on TLFB interview.

Analysis

We describe demographic characteristics for the VACS participants who completed the CIDI-SAM. Using the 3 reference measures, we calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), percent agreement, and kappa statistics for measures of alcohol use for both HIV-infected and -uninfected participants. We compared the following alcohol measures to the reference standard for risky drinking and unhealthy alcohol use: AUDIT, AUDIT-C, heavy episodic drinking, AUDIT or heavy episodic drinking, and AUDIT-C or heavy episodic drinking. For the unhealthy alcohol use reference standard, we also compared ICD-9 diagnosis of alcohol abuse or dependence in the past year. We compared those measures as well as the following measures to the reference standard for dependence: ICD-9 codes for alcohol diagnoses and an AUDIT dependence score (from items 4 to 6 of the AUDIT). The AUDIT and AUDIT-C cutoffs tested are based on prior research (Babor et al., 1989; Bradley et al., 1998b). Analyses were carried out using Stata 10.0 (College Station, TX).

RESULTS

Subject Characteristics

All subjects were men. Age and percentage with HIV were similar between the substudy sample (n = 837) and the entire VACS cohort (n = 6,467), although those in the analytic sample were more likely to be White than overall VACS participants (33 vs. 22%). Of the 837 in the analytic sample, half were HIV-infected, which is consistent with the study design target to recruit a similar number of HIV-infected and -uninfected patients. Race/ethnicity was similar between the HIV-infected and -uninfected participants with 53% African American, 8% Hispanic, 33% White, and 5% other. However, compared with uninfected subjects, HIV-infected were younger (average age, 50 vs. 54 years).

On the basis of the TLFB, 13% reported risky drinking amounts in the past 30 days. Seven percent of subjects met criteria for past year alcohol dependence based on the CIDI-SAM. On the basis of a combination of TLFB-determined risky amounts and CIDI-SAM-determined alcohol use disorders, 21% had unhealthy alcohol use. The prevalence of various levels of unhealthy use and alcohol measures is shown in Table 1, and there are no statistically significant differences in these alcohol measures between HIV-infected and -uninfected subjects.

Table 1. Subject Characteristics

	Total (n = 837)	HIV+ (n = 444)	HIV− (n = 393)	p-Value
Race/ethnicity (%)				
African American	53	56	51	0.5
Hispanic	8	7	8	
White	33	32	36	
Other	5	5	5	
Mean age, in years	52 (9.3)	50 (8.4)	54 (9.8)	<0.001
Alcohol use (%)				
Unhealthy use in past year (CIDI-SAM and TLFB)	21	22	20	0.5
Alcohol dependence in past year	7	8	7	0.6
Alcohol abuse/ dependence in past year	16	16	16	0.9
Risky consumption (TLFB)	13	14	12	0.8
AUDIT 8+	12	12	11	0.7
AUDIT 8+ or HED	26	25	27	0.7
AUDIT-C 6+	10	9	11	0.3
AUDIT-C 5+	15	15	15	0.8
AUDIT-C 4+	22	21	22	0.8
AUDIT-C 3+	31	30	31	0.8
AUDIT-C 6+ or HED	24	23	26	0.4
AUDIT-C 5+ or HED	25	24	26	0.4
AUDIT-C 4+ or HED	28	27	29	0.4
AUDIT-C 3+ or HED	33	32	34	0.5
HED ever	24	23	26	0.4
HED less than monthly	12	11	13	0.4
HED at least monthly	6	5	8	0.1
Alcohol-related ICD-9 diagnosis in past year	9	9	9	0.9
ICD-9 alcohol dependence diagnosis in past year	6	5	6	0.7
AUDIT 20+	3	2	3	0.6
AUDIT 16+	4	4	4	0.9
AUDIT dependence score 4+ (items 4 to 6)	4	3	4	0.7

CIDI-SAM, Composite International Diagnostic Interview-Substance Abuse Module; HED, heavy episodic drinking; AUDIT, Alcohol Use Disorders Identification Test; ICD-9; International Classification of Diseases, 9th revision; TLFB, Timeline Follow Back.

Alcohol Screening Measures and Risky Drinking

For HIV-infected and -uninfected subjects, sensitivity was 80% or greater between the TLFB reference standard for risky drinking and the following measures and cutoffs: AUDIT 4+ (80 vs. 86%), AUDIT 5+ or heavy episodic drinking (83 vs. 84%), AUDIT 4+ or heavy episodic drinking (83 vs. 90%) AUDIT-C 3+ (83 vs. 90%), AUDIT-C 4+ or heavy episodic drinking (83 vs. 88%), and AUDIT-C 3+ or heavy episodic drinking (85 vs. 94%). Specificity was 90% or greater for AUDIT 8+ (92 vs. 94%), AUDIT 7+ (90 vs. 92%), AUDIT-C 6+ (96 vs. 95%), AUDIT-C 5+ (92 vs. 94%), heavy episodic drinking less than monthly (94 vs. 94%), and at least monthly heavy episodic drinking (98 vs. 97%) (Table 2). The kappa statistics for these comparisons ranged from 0.28 to 0.61, indicating fair to moderate agreement (Landis and Koch, 1977).

Alcohol Screening Measures and Alcohol Dependence

For HIV-infected and -uninfected subjects, sensitivity was 80% or greater between the CIDI-SAM reference standard, for alcohol dependence in past year, and the following measures and cutoffs: AUDIT 4+ (83 vs. 81%), AUDIT 8+ or heavy episodic drinking (83 vs. 78%), AUDIT 7+ or heavy episodic drinking (83 vs. 81%), AUDIT 6+ or heavy episodic

Table 2. Agreement of "Gold Standard" for *Risky Drinking* Based on TLFB with Various Measures and Cutoffs

Comparison measure	Sens (%)	Spec (%)	PPV (%)	NPV (%)	Agree (%)	Kappa
HIV-infected						
AUDIT 8+	38	92	42	90	84	0.31
AUDIT 7+	52	90	44	92	85	0.39
AUDIT 6+	60	88	43	93	84	0.41
AUDIT 5+	73	84	42	95	82	0.43
AUDIT 4+	80	79	37	37	79	0.39
AUDIT 8+ or HED	75	82	40	95	81	0.42
AUDIT 7+ or HED	75	82	39	95	81	0.41
AUDIT 6+ or HED	78	81	39	96	81	0.42
AUDIT 5+ or HED	83	79	39	97	80	0.42
AUDIT 4+ or HED	83	75	34	97	76	0.37
AUDIT-C 6+	38	96	59	91	88	0.40
AUDIT-C 5+	60	92	54	94	88	0.49
AUDIT-C 4+	75	87	47	96	85	0.50
AUDIT-C 3+	83	78	37	97	79	0.40
AUDIT-C 2+	88	65	28	97	68	0.28
AUDIT-C 6+ or HED	72	84	41	95	83	0.43
AUDIT-C 5+ or HED	77	84	42	96	83	0.45
AUDIT-C 4+ or HED	80	82	41	96	82	0.44
AUDIT-C 3+ or HED	85	77	36	97	78	0.39
HED ever	72	85	42	95	83	0.43
HED less than monthly	42	94	52	91	87	0.39
HED at least monthly	23	98	61	89	88	0.28
HIV-uninfected						
AUDIT 8+	47	94	53	92	88	0.43
AUDIT 7+	55	92	51	93	87	0.45
AUDIT 6+	65	90	49	94	87	0.48
AUDIT 5+	75	88	48	96	86	0.50
AUDIT 4+	86	81	40	98	82	0.45
AUDIT 8+ or HED	82	82	40	97	82	0.44
AUDIT 7+ or HED	84	82	41	97	82	0.45
AUDIT 6+ or HED	84	82	41	97	82	0.45
AUDIT 5+ or HED	84	81	40	97	81	0.44
AUDIT 4+ or HED	90	77	37	98	79	0.42
AUDIT-C 6+	51	95	60	93	89	0.49
AUDIT-C 5+	71	94	63	96	91	0.61
AUDIT-C 4+	80	87	47	97	86	0.51
AUDIT-C 3+	90	77	37	98	79	0.42
AUDIT-C 2+	96	67	30	99	70	0.32
AUDIT-C 6+ or HED	82	82	41	97	82	0.45
AUDIT-C 5+ or HED	86	82	42	98	83	0.48
AUDIT-C 4+ or HED	88	80	39	98	81	0.45
AUDIT-C 3+ or HED	94	75	36	99	78	0.41
HED ever	81	83	40	97	83	0.45
HED less than monthly	56	94	56	94	89	0.50
HED at least monthly	37	97	61	92	89	0.40

AUDIT, Alcohol Use Disorders Identification Test; HED, heavy episodic drinking; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; Agree, agreement; TLFB, Timeline Follow Back.

drinking (86 vs. 81%), AUDIT 5+ or heavy episodic drinking (86 vs. 81%), and AUDIT-C 3+ or heavy episodic drinking (77 vs. 81%). Specificity was 90% or greater for the following measures: alcohol-related ICD-9 diagnosis in past year (93 vs. 92%), ICD-9 diagnosis code for dependence in the past year (97 vs. 94%), AUDIT 20+ (99 vs. 98%), AUDIT 16+ (98 vs. 98%), AUDIT 8+ (92 vs. 91%), AUDIT-C 10+ (100 vs. 99%), AUDIT-C 8+ (98 vs. 95%), AUDIT-C 6+ (94 vs. 91%), heavy episodic drinking less than monthly (92 vs. 90%) heavy episodic drinking at least monthly (98 vs. 94%), and AUDIT Dependence Score 4+ (99 vs. 98%) (Table 3). Among HIV-uninfected, the kappa statistics are 0.06 and 0.10 for the alcohol-related ICD-9 diagnosis and dependence alone, respectively. Otherwise, kappa statistics ranged from 0.17 to 0.43 indicating slight to moderate agreement (Landis and Koch, 1977).

Alcohol Screening Measures and Unhealthy Alcohol Use

For HIV-infected and -uninfected subjects, sensitivity was 80% or greater between the reference standard for unhealthy alcohol use and the following measures and cutoffs: AUDIT-C 2+ (86 vs. 96%) and AUDIT-C 3+ or heavy episodic drinking (71 vs. 83%). Specificity was 90% or greater for AUDIT 8+ (95 vs. 96%), AUDIT 7+ (93 vs. 95%), AUDIT 6+ (91 vs. 93%), AUDIT-C 6+ (97 vs. 97%), AUDIT-C 5+ (94 vs. 96%), heavy episodic drinking less than monthly (96 vs. 95%), at least monthly heavy episodic drinking (99 vs. 98%), and alcohol-related ICD-9 diagnosis (94 vs. 92%) (Table 4). For the alcohol-related ICD-9 diagnosis, kappa statistics are 0.16 and 0.02 for HIV-infected and -uninfected respectively. Otherwise, kappa statistics range from 0.27 to 0.57 indicating fair to moderate agreement (Landis and Koch, 1977).

DISCUSSION

In this sample, the prevalence of various types of unhealthy alcohol use were similar among HIV-infected and -uninfected adults. Compared with the reference standards, the AUDIT-C can identify risky drinking and unhealthy alcohol use with over 80% sensitivity at the best-performing cutoffs among both HIV-infected and -uninfected patients. Using the AUDIT-C combined with any heavy episodic drinking in the past year (item 3 of the AUDIT-C) improves sensitivity for detecting both risky and unhealthy alcohol use compared with using the AUDIT-C alone. High specificity for risky drinking and unhealthy alcohol use can be obtained using various AUDIT-C cutoffs and even using the heavy episodic drinking single-item question alone. We note that sensitivity is higher for most cutoffs for risky drinking than for unhealthy alcohol use. We believe this is because alcohol abuse and dependence are included in the definition for unhealthy alcohol use, and they may be more difficult to detect with AUDIT-based measures than risky drinking alone. For both risky drinking and unhealthy alcohol use, many cutoffs provide moderate agreement or better based on the kappa statistics.

For identifying those with alcohol dependence, the AUDIT performs better than the AUDIT-C and combining heavy episodic drinking with the AUDIT cutoffs results in higher sensitivity. Using a cutoff of AUDIT 6+ or heavy episodic drinking resulted in sensitivity of 86% for HIV-infected and 81% for HIV-uninfected. High specificity for alcohol dependence can be obtained using high cutoffs for the AUDIT (20+ and 16+), the AUDIT-C (10+ and 8+), the AUDIT dependence score of 4+, and the ICD-9 code for dependence in the past year; however, sensitivity is low (14 to

Table 3. Agreement of "Gold Standard" for *Alcohol Dependence* Based on CIDI-SAM with Various Measures and Cutoffs

Comparison measure	Sens (%)	Spec (%)	PPV (%)	NPV (%)	Agree (%)	Kappa
HIV-infected						
Alcohol-related ICD-9 diagnosis in past year	31	93	28	94	88	0.23
ICD-9 diagnosis of alcohol dependence in past year	26	97	39	94	91	0.26
AUDIT 20+	23	99	73	94	93	0.32
AUDIT 16+	31	98	58	94	93	0.37
AUDIT 8+	63	92	40	97	90	0.43
AUDIT 7+	69	89	34	97	87	0.39
AUDIT 6+	74	86	31	98	85	0.37
AUDIT 5+	74	80	25	97	80	0.28
AUDIT 4+	83	75	22	98	76	0.26
AUDIT 8+ or HED	83	79	26	98	80	0.31
AUDIT 7+ or HED	83	79	25	98	80	0.31
AUDIT 6+ or HED	86	78	25	98	79	0.30
AUDIT 5+ or HED	86	76	23	98	77	0.28
AUDIT 4+ or HED	86	72	21	98	73	0.24
AUDIT-C 10+	14	100	83	93	93	0.23
AUDIT-C 8+	29	98	56	94	93	0.34
AUDIT-C 6+	37	94	29	92	89	0.29
AUDIT-C 5+	57	88	30	96	86	0.32
AUDIT-C 4+	71	83	26	97	82	0.30
AUDIT-C 3+	74	74	19	97	74	0.21
AUDIT-C 2+	82	69	43	93	72	0.38
AUDIT-C 6+ or HED	74	81	25	97	81	0.29
AUDIT-C 5+ or HED	74	80	24	97	80	0.28
AUDIT-C 4+ or HED	77	78	23	97	78	0.26
AUDIT-C 3+ or HED	77	72	19	97	73	0.21
HED ever	74	81	25	97	81	0.30
HED less than monthly	46	92	33	95	89	0.32
HED at least monthly	37	98	57	95	93	0.41
AUDIT dependence score 4+ (items 4 to 6)	26	99	60	94	93	0.33
HIV-uninfected						
Alcohol-related ICD-9 diagnosis in past year	15	92	12	94	87	0.06
ICD-9 diagnosis of alcohol dependence in past year	15	94	17	94	89	0.10
AUDIT 20+	22	98	50	94	93	0.28
AUDIT 16+	33	98	56	95	94	0.39
AUDIT 8+	48	91	29	96	88	0.30
AUDIT 7+	52	89	25	96	86	0.27
AUDIT 6+	59	86	24	97	84	0.26
AUDIT 5+	74	84	25	98	83	0.30
AUDIT 4+	81	76	20	98	77	0.24
AUDIT 8+ or HED	78	77	20	98	77	0.23
AUDIT 7+ or HED	81	77	21	98	77	0.25
AUDIT 6+ or HED	81	77	21	98	77	0.25
AUDIT 5+ or HED	81	77	20	98	77	0.24
AUDIT 4+ or HED	81	72	18	98	73	0.20
AUDIT-C 10+	14	99	44	94	93	0.19
AUDIT-C 8+	26	95	29	95	91	0.22
AUDIT-C 6+	41	91	26	95	88	0.25
AUDIT-C 5+	59	89	28	97	87	0.32
AUDIT-C 4+	67	81	21	97	80	0.24
AUDIT-C 3+	74	72	16	97	72	0.17
AUDIT-C 2+	88	70	43	96	74	0.42
AUDIT-C 6+ or HED	74	78	20	98	77	0.23
AUDIT-C 5+ or HED	78	77	20	98	77	0.24
AUDIT-C 4+ or HED	78	75	18	98	75	0.21
AUDIT-C 3+ or HED	81	70	17	98	70	0.18
HED ever	74	78	20	98	78	0.23
HED less than monthly	52	90	28	96	88	0.30
HED at least monthly	26	94	23	94	89	0.19
AUDIT dependence score 4+ (items 4 to 6)	26	98	47	95	93	0.29

CIDI-SAM, Composite International Diagnostic Interview-Substance Abuse Module; ICD-9; International Classification of Diseases, 9th revision; AUDIT, Alcohol Use Disorders Identification Test; HED, heavy episodic drinking; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; Agree, agreement.

31%). Many cutoffs provide fair agreement or better based on the kappa statistics.

Our findings are consistent with other studies that have reported on recommendations for AUDIT and AUDIT-C cutoffs for unhealthy alcohol use. The standard cutoffs recommended for AUDIT range from 5 or more to 8 or more, and those recommended for the AUDIT-C range from 3 or more to 6 or more for various types of unhealthy alcohol use (Bradley et al., 1998b; Bush et al., 1998; Gordon et al., 2001; Gual et al., 2002). Additionally, our finding that the AUDIT-C performs as well as the longer AUDIT for identifying those with risky drinking and unhealthy alcohol use is consistent with findings of other researchers (Bradley et al., 1998b, 2009; Gordon and Saitz, 2004; Gordon et al., 2001). We demonstrate that adding a heavy episodic drinking component improves sensitivity compared with AUDIT cutoffs alone without adding additional burden to the patients or survey respondents as this question is already included as part of the AUDIT.

We found that the prevalence of risky drinking is slightly higher in the HIV-infected subjects based on the TLFB, but the percentage of patients identified with risky drinking is slightly lower among the HIV-infected subjects based on the various cutoffs of AUDIT, AUDIT-C, and heavy episodic drinking. For risky drinking and unhealthy alcohol use, we are uncertain as to why sensitivity is lower among the HIV-infected compared with uninfected. One study reported that those with HIV used alcohol at lower usual levels but with greater variability compared with those who were HIV-uninfected (Braithwaite et al., 2008). Perhaps this difference in drinking patterns between HIV-infected and -uninfected provides some insight into the differences in sensitivity. More likely, however, there may be a social desirability bias causing the HIV-infected patients to underreport their alcohol consumption.

For alcohol dependence, agreement is particularly low for the ICD-9 codes for the HIV-uninfected (kappa $\leq$ 0.10) compared with the HIV-infected subjects (kappa = 0.23 to 0.26). Although we realize that alcohol-related diagnoses often are underreported, we are surprised to see a lower level of agreement among the HIV-uninfected subjects given that the prevalence of all alcohol measures including the ICD-9 measures are similar between the HIV-infected

Table 4. Agreement of "Gold Standard" for *Unhealthy Alcohol Use* Based on CIDI-SAM and TLFB with Various Measures and Cutoffs

Comparison measure	Sens (%)	Spec (%)	PPV (%)	NPV (%)	Agree (%)	Kappa
HIV-infected						
AUDIT 8+	38	95	67	84	82	0.38
AUDIT 7+	47	93	66	86	83	0.45
AUDIT 6+	54	91	64	88	83	0.48
AUDIT 5+	63	87	58	89	82	0.49
AUDIT 4+	69	82	52	90	79	0.46
AUDIT 8+ or HED	66	86	58	90	82	0.50
AUDIT 7+ or HED	66	86	57	90	82	0.49
AUDIT 6+ or HED	69	85	57	91	82	0.50
AUDIT 5+ or HED	72	83	55	91	81	0.50
AUDIT 4+ or HED	73	79	50	91	78	0.45
AUDIT-C 6+	30	97	74	83	82	0.34
AUDIT-C 5+	49	94	71	87	84	0.49
AUDIT-C 4+	61	90	63	89	84	0.52
AUDIT-C 3+	70	81	51	91	79	0.45
AUDIT-C 2+	86	65	28	97	68	0.28
AUDIT-C 6+ or HED	61	88	58	89	82	0.48
AUDIT-C 5+ or HED	64	87	59	90	82	0.50
AUDIT-C 4+ or HED	66	85	55	90	81	0.48
AUDIT-C 3+ or HED	71	79	50	91	78	0.44
HED ever	61	88	59	89	82	0.48
HED less than monthly	36	96	73	84	83	0.39
HED at least monthly	20	99	87	81	82	0.27
Alcohol-related ICD-9 diagnosis in past year	18	94	46	80	77	0.16
HIV-uninfected						
AUDIT 8+	43	96	76	87	86	0.47
AUDIT 7+	48	95	69	88	85	0.48
AUDIT 6+	56	93	66	89	85	0.52
AUDIT 5+	65	91	65	91	86	0.56
AUDIT 4+	74	84	54	93	82	0.51
AUDIT 8+ or HED	74	85	56	93	83	0.53
AUDIT 7+ or HED	75	85	57	93	83	0.54
AUDIT 6+ or HED	75	85	57	93	83	0.54
AUDIT 5+ or HED	75	85	56	93	83	0.53
AUDIT 4+ or HED	79	81	51	94	80	0.49
AUDIT-C 6+	41	97	77	87	86	0.46
AUDIT-C 5+	55	96	77	89	88	0.57
AUDIT-C 4+	65	89	60	91	84	0.52
AUDIT-C 3+	76	80	50	93	79	0.47
AUDIT-C 2+	96	67	30	99	70	0.32
AUDIT-C 6+ or HED	71	86	56	92	83	0.52
AUDIT-C 5+ or HED	74	86	57	93	83	0.54
AUDIT-C 4+ or HED	76	83	54	93	82	0.51
AUDIT-C 3+ or HED	83	79	50	95	79	0.49
HED Ever	70	86	55	92	82	0.51
HED less than monthly	44	95	70	87	85	0.45
HED at least monthly	29	98	77	84	84	0.35
Alcohol-related ICD-9 diagnosis in past year	10	92	24	80	75	0.02

CIDI-SAM, Composite International Diagnostic Interview-Substance Abuse Module; ICD-9; International Classification of Diseases, 9th revision; AUDIT, Alcohol Use Disorders Identification Test; HED, heavy episodic drinking; Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; Agree, agreement; TLFB, Timeline Follow Back.

and -uninfected groups. Perhaps HIV-infected patients are screened more frequently or aggressively because of increased surveillance for the potential negative interactions of alcohol use with medications and/or other comorbidities, such as hepatitis C.

This study has several strengths. It uses a large U.S. multisite cohort of racially/ethnically diverse HIV-infected and -uninfected patients with multiple measures of alcohol use. Many studies have evaluated and identified the optimal cutoffs to use for the AUDIT and the AUDIT-C to identify risky drinking amounts, alcohol dependence, and the spectrum of unhealthy alcohol use. However, no other studies that we are aware of have evaluated these measures in an HIV-infected population. This study provides new information regarding the performance of commonly used screening and diagnostic instruments among HIV-infected subjects, an important population that is at higher risk for adverse outcomes from alcohol use. For many studies and clinical settings, it is important to have brief tools for evaluating alcohol use to help minimize participant burden and maximize time and resources, especially when one is attempting to collect information on a variety of behaviors. In clinical settings especially, patient assessments need to be brief because of patient and provider time constraints.

This study has several limitations. First, our study sample was restricted to Veterans receiving medical care and, because the majority of our cohort is men, we limited the analysis to men. Similar studies should be conducted on cohorts of HIV-infected and -uninfected women. The mean age of our study population is 52 years, and this should be considered when applying these findings to older or younger populations. For risky drinking, we compared the AUDIT, which captures data for the previous year, with the TLFB, which captures drinking information over the previous 30 days. On the basis of these differing time frames, we might expect to see lower prevalence of risky drinking identified by the TLFB, especially if we are incorporating heavy episodic drinking, which we expect to have reported more frequently over the past year in the AUDIT than in the past 30 days on the TLFB. How much prevalence differs would depend on how consistent drinking behavior was over time; furthermore, the AUDIT asks about typical drinking that is likely affected by recent (e.g., past month) behavior. In any case, sensitivity should not be impacted by the time frame discrepancy between measures because the screening test covers a longer time than the reference standard. However, specificity could be lower because of some individuals not meeting criteria for risky drinking over the past 30 days on the TLFB, but meeting risky drinking criteria based on AUDIT or AUDIT-C during the previous year.

Although our results demonstrate that agreement is similar between HIV-infected and -uninfected subjects for identifying participants on the spectrum of alcohol use, we cannot be sure that screening tests administered in person as 1 to 3 items as part of clinical care will perform the same as the screening tests asked as part of a research survey that is self-administered.

Despite these limitations, our findings have implications for the use of these screening measures in clinical practice and research. We cannot recommend a standard instrument or cutoff to use for all purposes, as the decision will depend on the purpose (e.g., screening) for each situation. However,

based on these findings, the optimal balance between sensitivity, specificity, and agreement is provided by the following instruments using the specified cutoffs for both general medical and HIV care settings: AUDIT-C 4+ or heavy episodic drinking for risky drinking; AUDIT 6+ or heavy episodic drinking for alcohol dependence; and AUDIT-C 3+ or heavy episodic for unhealthy alcohol use. Of note, for unhealthy alcohol use, few cutoffs provide sensitivity $\geq 80\%$; however, many cutoffs provide good specificity. For identifying alcohol dependence, the highly specific AUDIT dependence score and the ICD-9 diagnosis codes are not recommended when sensitivity is important, such as for screening measures by physicians. However, in the absence of other sources of alcohol data, ICD-9 codes can be useful where specificity is important and sensitivity is not (e.g., when evaluating costly and/or time-intensive interventions).

In summary, for both HIV-infected and -uninfected patients, briefer measures of alcohol use provide reasonable sensitivity and specificity for risky and unhealthy use compared with the longer more resource intensive measures. The incorporation of heavy episodic drinking from question 3 of the AUDIT with AUDIT and AUDIT-C cutoffs improves sensitivity for identifying risky drinking, alcohol dependence, and unhealthy alcohol use compared with using AUDIT and AUDIT-C cutoffs alone. For many research and clinical purposes, the briefer measures should be sufficient, at least for screening purposes, and then more time-intensive measures could be utilized to follow-up on positive screens. Cutoffs can be chosen to maximize sensitivity or specificity as appropriate to the clinical or research application. For alcohol dependence, the briefer measures may not be adequate for more than broad screening; a more intensive measure, such as the CIDI-SAM, may still be necessary when sensitivity is important. However, when specificity is important and sensitivity is not, brief tools can be used when identifying people with alcohol dependence.

ACKNOWLEDGMENT

The Veterans Aging Cohort Study is funded by the National Institute on Alcohol Abuse and Alcoholism (U10 AA 13566).

REFERENCES

Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG (1989) AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. 2nd ed. Department of Mental Health and Substance Dependence, World Health Organization, Geneva, Switzerland.

Baum MK, Rafie C, Lai S, Sales S, Page JB, Campa A (2010) Alcohol use accelerates HIV disease progression. AIDS Res Hum Retroviruses 26: 511–518.

Bradley KA, Boyd-Wickizer J, Powell SH, Burman ML (1998a) Alcohol screening questionnaires in women: a critical review. JAMA 280:166–171.

Bradley KA, Kivlahan DR, Williams EC (2009) Brief approaches to alcohol screening: practical alternatives for primary care. J Gen Intern Med 24: 881–883.

Bradley KA, McDonell MB, Bush K, Kivlahan DR, Diehr P, Fihn SD (1998b) The AUDIT alcohol consumption questions: reliability, validity, and responsiveness to change in older male primary care patients. Alcohol Clin Exp Res 22:1842–1849.

Braithwaite RS, Bryant K (2010) Influence of alcohol consumption on adherence to and toxicity of antiretroviral therapy and survival. Alcohol Res Health 33:280–297.

Braithwaite RS, Conigliaro J, McGinnis KA, Maisto SA, Bryant K, Justice AC (2008) Adjusting alcohol quantity for mean consumption and intoxication threshold improves prediction of nonadherence in HIV patients and HIV-negative controls. Alcohol Clin Exp Res 32:1645–1651.

Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Arch Intern Med 158:1789–1795.

Compton WM (1993) Advantages of the CIDI family of instruments in epidemiological research of substance use disorders. Int J Methods Psychiatr Res. Special Issue: The WHO Composite International Diagnostic Interview 3:109–119.

Conigliaro J, Gordon AJ, McGinnis KA, Rabeneck L, Justice AC (2003) How harmful is hazardous alcohol use and abuse in HIV infection: do health care providers know who is at risk? J Acquir Immune Defic Syndr 33:521–525.

Dawson DA, Grant BF, Stinson FS, Zhou Y (2005) Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the U.S. general population. Alcohol Clin Exp Res 29:844–854.

Fleming MF (2004/2005) Screening and brief intervention in primary care settings. Alcohol Res Health 28:57–62.

Gordon AJ, Maisto S, McNeil M, Kraemer KL, Conigliaro RL, Kelley ME, Conigliaro J (2001) Three questions can detect hazardous drinkers. J Fam Pract 50:313–320.

Gordon AJ, Saitz R (2004) Identification and management of alcohol use disorders in primary care. J Clin Outcomes Manag 11:444–460.

Gual A, Segura L, Contel M, Heather N, Colom J (2002) AUDIT-3 and AUDIT-4: effectiveness of two short forms of the alcohol use disorders identification test. Alcohol Alcohol 37:591–596.

Hansagi H, Romelsjo A, Gerhardsson de Verdier M, Andreasson S, Leifman A (1995) Alcohol consumption and stroke mortality. 20-year follow-up of 15,077 men and women. Stroke 26:1768–1773.

Justice AC, Dombrowski E, Conigliaro J, Fultz SL, Gibson D, Madenwald T, Goulet J, Simberkoff M, Butt AA, Rimland D, Rodriguez-Barradas MC, Gibert CL, Oursler KA, Brown S, Leaf DA, Goetz MB, Bryant K (2006a) Veterans Aging Cohort Study (VACS): overview and description. Med Care 44:S13–S24.

Justice AC, Lasky E, McGinnis KA, Skanderson M, Conigliaro J, Fultz SL, Crothers K, Rabeneck L, Rodriguez-Barradas M, Weissman SB, Bryant K (2006b) Medical disease and alcohol use among veterans with human immunodeficiency infection: a comparison of disease measurement strategies. Med Care 44:S52–S60.

Kessler RC, Wittchen HU, Abelson JM, Kendler KS, Knauper B, Zhao S (1998) Methodological studies of the composite international diagnostic interview (CIDI) in the US national comorbidity survey. Int J Methods Psychiatr Res 7:33–55.

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174.

Maisto SA, Gordon A, Conigliaro JC, McGinnis KA, Justice AC (2009) An experimental study of the agreement of self-and telephone administration of the Timeline Followback Interview. J Stud Alcohol Drugs 69: 468–471.

Naimi TS, Brewer RD, Mokdad A, Denny C, Serdula MK, Marks JS (2003) Binge drinking among US adults. JAMA 289:70–75.

Piette J, Barnett PG, Moos R (1998) First-time admissions with alcohol-related medical problems: a 10-year follow-up of a national sample of alcoholic patients. J Stud Alcohol 59:89–96.

Raj A, Reed E, Santana MC, Walley AY, Welles SL, Horsburgh CR, Flores SA, Silverman JG (2009) The associations of binge alcohol use with HIV/

STI risk and diagnosis among heterosexual African American men. Drug Alcohol Depend 101:101–106.

Reinert DF, Allen JP (2002) The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. Alcohol Clin Exp Res 26:272–279.

Rubinsky AD, Kivlahan DR, Volk RJ, Maynard C, Bradley KA (2010) Estimating risk of alcohol dependence using alcohol screening scores. Drug Alcohol Depend 108:29–36.

Saitz R (2005) Clinical practice. Unhealthy alcohol use. N Engl J Med 352:596–607.

Samet JH, Horton NJ, Traphagen ET, Lyon SM, Freedberg KA (2006) Alcohol consumption and HIV disease progression: are they related? Alcohol Clin Exp Res 27:862–867.

Saunders JB, Aasland OG, Babor TF, DeLaFuente JR, Grant M (1993) Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. Addiction 88:791–804.

Schmidt A, Barry KL, Fleming MF (1995) Detection of problem drinkers: the alcohol use disorders identification test (AUDIT). South Med J 88:52–59.

Sobell L, Sobell M, Leo G, Cancilla A (1988a) Reliability of a timeline method: assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. Br J Addict 83:393–402.

Sobell L, Sobell M, Riley D, Schuller R, Pavan D, Cancilla A, Klainer F, Leo GI (1988b) The reliability of alcohol abusers' self-reports of drinking and life events that occurred in the distant past. J Stud Alcohol 49: 225–232.

Sobell LC, Sobell MB (1992) Timeline Followback: A technique for assessing self-reported ethanol consumption, in *Measuring Alcohol Consumption: Psychosocial and Biological Methods* (Allen J, Litten RZ eds), pp 41–72. Humana Press, Totowa, NJ.

Sullivan LE, Goulet JL, Justice AC, Fiellin DA (2011) Alcohol consumption and depressive symptoms over time: a longitudinal study of patients with and without HIV infection. Drug Alcohol Depend 117:158–163.

Journal of Health Communication, 17:7–12, 2012

ISSN: 1081-0730 print/1087-0415 online
DOI: 10.1080/10810730.2012.712627

Routledge
Taylor & Francis Group

Commentary

Health Literacy and the Disenfranchised: The Importance of Collaboration Between Limited English Proficiency and Health Literacy Researchers

MICHAEL M. MCKEE

Family Medicine Research Programs, Department of Family Medicine, University of Rochester School of Medicine and Dentistry, Rochester, New York, USA

MICHAEL K. PAASCHE-ORLOW

Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

Inadequate health literacy and limited English proficiency are associated with poor health care access and outcomes. Despite what appears to be an interaction phenomenon—whereby the rate of inadequate health literacy is particularly high among limited English proficiency populations—researchers in health literacy and limited English proficiency rarely collaborate. As a result, few health literacy instruments and interventions have been developed or validated for smaller linguistic populations. Interventions to improve health outcomes for people with low health literacy and limited English proficiency show great potential to alleviate many of the health disparities currently experienced by some of the most disenfranchised individuals in our health care system, those from smaller linguistic minority groups, including Deaf American Sign Language users. It is critical for health literacy and limited English proficiency researchers to work together to understand how culture, language, literacy, education, and disabilities influence health disparities and health outcomes. It is important to ensure that research is collaborative and inclusive in order to broaden the reach of future interventions to smaller linguistic minority populations.

The health care system is struggling to care for an increasingly diverse patient population (National Center for Education Statistics, 2001; Smedley, Stith, & Nelson, 2003; U.S. Census Bureau, 2011). It is unfortunate that the details of this diversity, including language preference, literacy, and culture, have in general been examined independently. Few studies have concurrently explored health literacy and limited English proficiency (LEP). A recent PubMed search by the coauthors using the health

Michael M. McKee is currently supported by grant K01 HL103140-01 from the National Heart, Lung, and Blood Institute of the National Institutes of Health.

Address correspondence to Michael M. McKee, Department of Family Medicine Research Programs, 1381 South Avenue, Rochester, NY 14620, USA. E-mail: michael_mckee@urmc.rochester.edu

literacy and LEP as MeSH keywords generated 5,158 health literacy references and 595 LEP references, but only 36 overlapping references. The paucity of publications on this topic underscores the need for better collaboration between these fields.

More than 90 million Americans have inadequate health literacy (Nielsen-Bohlman, Panzer, & Kindig, 2004) and more than 24.5 million Americans experience LEP (U.S. Census Bureau, 2010), not accounting for people with LEP among the population of unauthorized or undocumented immigrants in the United States (Passel & Cohn, 2011). In addition, this estimate does not include approximately 500,000 Deaf American Sign Language (ASL) users (Harrington, 2004; Schick, 1988).

It is not well established what proportion of LEP individuals have inadequate health literacy, but overall it is thought to be higher than that of fluent English speakers (Berkman et al., 2004; Nielsen-Bohlman et al., 2004). Several populations at risk for LEP (no English spoken before starting school, Hispanic adults, and those with disabilities) are overrepresented among those with below basic prose literacy on the National Assessment of Adult Literacy (Baldi, 2009). However, more important than establishing a precise prevalence estimate of inadequate health literacy and LEP is to understand their health consequences. In isolation, both pose significant barriers to health care communication (Baker, Parker, Williams, & Clark, 1998; Fernandez et al., 2004; Graham, Jacobs, Kwan-Gett, & Cover, 2008; Paasche-Orlow & Wolf, 2007; Schillinger et al., 2003; Wilson et al., 2005). Together, the damaging effects of inadequate health literacy and LEP on health communication and outcomes are likely magnified (Sudore et al., 2009). This may be due to synergistic negative effects in phenomena such as socioeconomic position, access to care, access to information, as well as perceived and experienced discrimination (Egede, 2006).

Quality measures mandated by the Centers for Medicare and Medicaid Services heavily rely on improved patient communication and outcomes (Centers for Medicare & Medicaid Services, 2012). There is great interest among hospitals, managed care, and outpatient health centers to develop effective communication strategies to deal with diverse patient populations. Many of these approaches will require interventions to deal with the complexities of handling patient populations who experience LEP and inadequate health literacy. Unfortunately, the current research environment is poorly equipped to provide innovative approaches and interventions to increase the ability of a health consumer with LEP and inadequate health literacy to make appropriate health care decisions. Researchers must work together to understand how culture, language, literacy, education, and disabilities all play elemental roles in promulgating health disparities and health outcomes.

Benefits of Collaborative Research

Despite an increasing number of interventions designed to improve health knowledge, and disease management, few studies evaluate whether these interventions are reproducible in smaller language minority populations and among individuals with a range of health literacy (Sarkar et al., 2008; Schillinger et al., 2008). Many interventions intended to address LEP may also be appropriate for those with low health literacy as well as the converse. The hybridization of LEP and health literacy research would further enhance the ability to develop novel approaches and interventions that can potentially be interchangeable for diverse communication needs. Furthermore, health literacy research could greatly benefit from the infusion of transdisciplinary approaches provided in the fields of linguistics and cognitive science. LEP research

could gain from standardized measures and instruments available and emerging in health literacy research.

LEP research has demonstrated the benefits of language concordance and the use of interpreters including higher rates of appropriate health care visits, improved treatment adherence, and improved satisfaction (Graham et al., 2008; MacKinney, Walters, Bird, & Nattinger, 1995; McKee, Barnett, Block, & Pearson, 2011; Regenstein et al., 2008; Timmins, 2002). Interventions involving the use of community health workers, health coaches, language interpreters, and language-concordant providers are common approaches for LEP populations. However, it is unclear whether these interventions can improve health literacy among these populations. In addition, novel approaches for inadequate health literacy through the use of technology (e.g., virtual health coaches) have not yet been adequately examined in LEP populations. Health information technologies have promising potential to tailor health information to language and cultural needs and help people access to health materials in a variety of modalities when and where they may be needed (Munoz et al., 2009). Further research is needed to determine how this can be applied effectively for people with LEP and inadequate health literacy.

Populations Who May Benefit From Collaborative Research

Spanish-speaking populations, by virtue of their size, have garnered the majority of LEP and non–English health literacy research funding in the United States. Despite the fact that more than 376 languages are used in the United States, a review of the literature reveals very few of these studies extend beyond English and Spanish. Schillinger, Bindman, Wang, Stewart, and Piet (2004) demonstrated that even when accounting for language barriers, inadequate reading and health literacy were highly prevalent in one Latino population sample; only 55% of Hispanics have a high school diploma, and only 10% have a bachelor's degree (Schillinger et al., 2004). Inadequate health literacy may be even higher for smaller linguistic minority groups as a result of a paucity of accessible media and patient education.

Certain immigrant populations and Deaf ASL users live in relative isolation. These populations provide unique research challenges and opportunities because of reduced social interactions with the media and limited contact with allopathic norms, public health and prevention messaging, and health education. These groups experience poor patient–provider communication and frequently rely on inaccurate and inconsistent information from their social networks and the Internet (McKee et al., 2011; Valentine & Skelton, 2009; Vernon & Lynch, 2003).

Deaf ASL users, in particular, present a unique study population because they struggle with poor communication (i.e., due to hearing loss), language discordance, and possibly inadequate health literacy partly because of decreased opportunities to correct misinformation, and limited health surveillance (McKee et al., 2012). Deaf ASL users are considered an LEP population by the U.S. Department of Health and Human Services (U.S. Department of Health and Human Services, 2001); yet, they are rarely included in health literacy and LEP research studies. This is likely due to limited health and demographic information about Deaf ASL users, scarcity of adapted and validated research surveys and instruments in ASL, difficulty recruiting and accessing this population for research, inadequate supply of Deaf ASL proficient investigators, and concern about handling potential cognitive issues when present.

The exclusion of linguistic minorities is ultimately due to the fact that they are minorities. The research funding and the workforce are limited. Investigators want to have the largest possible impact and have research products that reach a critical mass. Although smaller language populations may be challenging to recruit and study, they are often most in need; work with such populations underscores a striking absence of health information accessible in languages beyond English and Spanish. Rudd and Anderson (2006) argued that existing health literacy interventions can be modified to apply to at-risk individuals and that environmental evaluations of health care systems to reduce literacy and language barriers could benefit a broad array of patients. The same approach to currently existing health literacy and LEP research tools and interventions can help identify critical steps needed to create greater inclusivity in research.

Future Implications

LEP and health literacy researchers should design research that advances knowledge regarding the intersection between these fields of inquiry. Approaches likely to promote this agenda include the following:

- Funding agencies should encourage collaboration between researchers through targeted requests for applications for research that incorporates both fields.
- Researchers should place special emphasis in developing tools and disseminating interventions that can be readily adapted and translated into languages for validation and use in other populations.
- Institutional review boards should give special consideration to smaller linguistic minority populations, including Deaf ASL users, to ensure that research is inclusive whenever possible.
- LEP and health literacy investigators should design research that implements novel strategies using information technology and other methods to lower communication barriers and bridge health care gaps more effectively and efficiently.
- Opportunities for collaboration and dissemination of innovative cross-disciplinary approaches, such as symposia and conferences with a special emphasis on bringing health literacy and LEP investigators together should be advanced.
- Academic institutions and research facilities should increase the number of researchers and staff from diverse backgrounds to improve recruitment of challenging-to-reach populations and provide greater social and cultural awareness necessary to establish rapport with targeted populations.

Conclusion

The fields of LEP and health literacy have largely functioned as separate silos of research, failing to address the needs of these unique and underserved populations in the U.S. health care system. It is clear that collaboration is needed between experts in these fields to help develop a variety of interventions and tools to assure the most vulnerable patients are not left behind. Without attention to the interaction between health literacy and LEP populations are at great risk for experiencing increasing health disparities.

Given that language and health literacy are integral to patients' ability to comprehend and act upon health recommendations, it is crucial to find ways to understand the joint effects of these phenomena. Researchers should view health literacy through

multicultural and multilingual lenses to help develop novel communication strategies and technologies that can be implemented in the increasing number of linguistic minority and LEP populations, including Deaf ASL users in the United States.

References

Baker, D. W., Parker, R. M., Williams, M. V., & Clark, W. S. (1998). Health literacy and the risk of hospital admission. *Journal of General Internal Medicine*, *13*, 791–798.

Baldi, S. (2009). *Technical Report and Data File User's Manual for the 2003 National Assessment of Adult Literacy (NCES 2009-476)*. Retrieved from http://nces.ed.gov/pubs2009/2009476.pdf

Berkman, N., DeWalt, D., Pignone, M., Sheridan, S., Lohr, K., Lux, L., … Bonito, A. (2004). *Literacy and health outcomes*. Rockville, MD: Agency for Health Care Research and Quality.

Centers for Medicare & Medicaid Services. (2012). *Quality initiatives–General information.* Retrieved from http://www.cms.gov/QualityInitiativesGenInfo/01_overview.asp

Egede, L. E. (2006). Race, ethnicity, culture, and disparities in health care. *Journal of General Internal Medicine*, *21*, 667–669.

Fernandez, A., Schillinger, D., Grumbach, K., Rosenthal, A., Stewart, A. L., Wang, F., & Pérez-Stable, E. J. (2004). Physician language ability and cultural competence. An exploratory study of communication with Spanish-speaking patients. *Journal of General Internal Medicine*, *19*, 167–174.

Graham, E. A., Jacobs, T. A., Kwan-Gett, T. S., & Cover, J. (2008). Health services utilization by low-income limited English proficient adults. *Journal of Immigrant & Minority Health*, *10*, 207–217.

Harrington, T. (2004). *American Sign Language: Ranking and number of "speakers."* Retrieved from http://libguides.gallaudet.edu/content.php?pid=114804&sid=991835

MacKinney, T. G., Walters, D., Bird, G. L., & Nattinger, A. B. (1995). Improvements in preventive care and communication for deaf patients: Results of a novel primary health care program. *Journal of General Internal Medicine*, *10*, 133–137.

McKee, M., Barnett, S., Block, R., & Pearson, T. (2011). Impact of communication on preventive services among deaf American Sign Language users. *American Journal of Preventive Medicine*, *41*, 75–79.

McKee, M., Schlehofer, D., Cuculick, J., Starr, M., Smith, S., & Chin, N. P. (2011). Perceptions of cardiovascular health in an underserved community of deaf adults using American Sign Language. *Disability & Health Journal*, *4*, 192–197.

McKee, M., Thew, D., Starr, M., Kushalnagar, P., Reid, J. T., Graybill, P., … Pearson, T. A. (2012). Engaging the deaf sign language community: Lessons from a community-based participatory research (CBPR) center. *Progress in Community Health Partnerships: Research, Education, and Action*, *6*, 319–328.

Munoz, R. F., Barrera, A. Z., Delucchi, K., Penilla, C., Torres, L. D., & Perez-Stable, E. J. (2009). International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. *Nicotine & Tobacco Research*, *11*, 1025–1034.

National Center for Education Statistics. (2001). *English literacy and language minorities in the United States*. Retrieved from http://nces.ed.gov/pubs2001/2001464.pdf

Nielsen-Bohlman, L., Panzer, A., & Kindig, D. (Eds.). (2004). *Health literacy: A prescription to end confusion*. Washington, DC: National Academies Press.

Paasche-Orlow, M. K., & Wolf, M. S. (2007). The causal pathways linking health literacy to health outcomes. *American Journal of Health Behavior, 31*(Suppl 1), S19–S26.

Passel, J. S., & Cohn, D. V. (2011, Febuary). *Unauthorized Immigrant Population: National and State Trends, 2010.* Retrieved from http://www.pewhispanic.org/files/reports/133.pdf

Regenstein, M., Huang, J., West, C., Mead, H., Trott, J., & Stegun, M. (2008). Hospital language services: Quality improvement and performance measures In K. Henriksen, J. B. Battles, M. A. Keyes, & M. L. Grady (Eds.), *Advances in Patient Safety: New Directions and Alternative Approaches* (Vol. 2). Retrieved from http://www.ahrq.gov/downloads/pub/advances2/vol2/Advances-Regenstein_54.pdf

Rudd, R., & Anderson, J. (2006). *The health literacy environment of hospitals and health centers: Partners for action: Making your healthcare facility literacy friendly.* Retrieved from http://www.hsph.harvard.edu/healthliteracy/files/healthliteracyenvironment.pdf

Sarkar, U., Piette, J. D., Gonzales, R., Lessler, D., Chew, L. D., Reilly, B., … Schillinger, D. (2008). Preferences for self-management support: findings from a survey of diabetes patients in safety-net health systems. *Patient Education & Counseling*, *70*, 102–110.

Schick, B. (1988). *American Sign Language: What is it? Who uses it?* Retrieved from http://spot.colorado.edu/~schick/ASL_info/asl_info.html

Schillinger, D., Bindman, A., Wang, F., Stewart, A., & Piette, J. (2004). Functional health literacy and the quality of physician–patient communication among diabetes patients. *Patient Education & Counseling*, *52*, 315–323.

Schillinger, D., Hammer, H., Wang, F., et al. (2008). Seeing in 3D: Examining the reach of diabetes selfmanagement support strategies in a public health care system. *Health Education and Behavior*, *35*, 664–682.

Schillinger, D., Piette, J., Grumbach, K., Wang, F., Wilson, C., Daher, C., … Bindman, A. B. (2003). Closing the loop: Physician communication with diabetic patients who have low health literacy. *Archives of Internal Medicine*, *163*, 83–90.

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (Eds.). (2003). *Unequal treatment: Confronting racial and ethnic disparities in health care (full printed version)*. Washington, DC: National Academies Press.

Sudore, R. L., Landefeld, C. S., Perez-Stable, E. J., Bibbins-Domingo, K., Williams, B. A., & Schillinger, D. (2009). Unraveling the relationship between literacy, language proficiency, and patient–physician communication. *Patient Education & Counseling*, *75*, 398–402.

Timmins, C. (2002). The impact of language barriers on the health care of Latinos in the United States: A review of the literature and guidelines for practice. *Journal of Midwifery & Women's Health*, *47*, 80–96.

U.S. Census Bureau. (2010). *Language use in the United States: 2007, American Community Survey reports*. Retrieved from http://www.census.gov/prod/2010pubs/acs-12.pdf

U.S. Census Bureau. (2011). *Overview of race and Hispanic origin: 2010*. Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf

U.S. Department of Health and Human Services. (2001). *National standards for culturally and linguistically appropriate services in health care*. Retrieved from http://minorityhealth.hhs.gov/assets/pdf/checked/finalreport.pdf

Valentine, G., & Skelton, T. (2009). "An umbilical cord to the world"—The role of the Internet in deaf people's information and communication practices. *Information, Communication & Society*, *12*, 44–65.

Vernon, R., & Lynch, D. (2003). Consumer access to agency websites: Our best foot forward? *Journal of Technology in Human Services*, *21*, 37–50.

Wilson, E., Chen, A., Grumbach, K., Wang, F., & Fernandez, A. (2005). Effects of limited English proficiency and physician language on health care comprehension. *Journal of General Internal Medicine*, *20*, 800–806.

Discrepancy in Diagnosis and Treatment of Post-traumatic Stress Disorder (PTSD): Treatment for the Wrong Reason

Ellen C. Meltzer, MD, MSc
Tali Averbuch, MPP
Jeffrey H. Samet, MD, MA, MPH
Richard Saitz, MD, MPH
Khelda Jabbar, MD
Christine Lloyd-Travaglini, MPH
Jane M. Liebschutz, MD, MPH

Abstract

In primary care (PC), patients with post-traumatic stress disorder (PTSD) are often undiagnosed. To determine variables associated with treatment, this cross-sectional study assessed 592 adult patients for PTSD. Electronic medical record (EMR) review of the prior 12 months assessed mental health (MH) diagnoses and MH treatments [selective serotonin reuptake inhibitor (SSRI) and/or ≥1 visit with MH professional]. Of 133 adults with PTSD, half (49%; 66/133) received an SSRI (18%), a visit with MH professional (14%), or both (17%). Of those treated, 88% (58/66) had an EMR MH diagnosis, the majority (71%; 47/66) depression and (18%; 12/66) PTSD. The odds of receiving MH treatment were increased 8.2 times (95% CI 3.1–21.5) for patients with an EMR MH diagnosis. Nearly 50% of patients with PTSD received MH treatment, yet few had this diagnosis documented. Treatment was likely due to overlap in the management of PTSD and other mental illnesses.

Address correspondence to Jane M. Liebschutz, MD, MPH, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA. Phone: +1-617-4147399; Fax: +1-617-4144676; Email: jane.liebschutz@bmc.org Email: jliebs@bu.edu.

Ellen C. Meltzer, MD, MSc, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA.

Tali Averbuch, MPP, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA.

Jeffrey H. Samet, MD, MA, MPH, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA.

Richard Saitz, MD, MPH, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine, 801 Massachusetts

Journal of Behavioral Health Services & Research, 2011.

Introduction

In primary care (PC) settings, patients with post-traumatic stress disorder (PTSD), are often not diagnosed; 2% to 11% with PTSD actually have the diagnosis noted in the medical record.[1,2] In addition, less than half of these patients with PTSD, or even fewer, actually receive treatment for PTSD.[3,4] To better address this condition in practice, more attention will likely need to be focused both on the recognition of PTSD and the treatment of PTSD when it is recognized.

Mental health (MH) treatment for PTSD includes pharmacotherapy and/or specialized MH counseling with structured cognitive behavioral therapy or psychotherapy.[5–9] Recent practice guidelines from the American Psychiatric Association recommend selective serotonin reuptake inhibitors (SSRIs) as first-line pharmacotherapy for patients with PTSD, given that they alleviate symptoms of PTSD, have few side effects, and also treat the comorbid depression, anxiety, and panic disorder that frequently co-occur with PTSD.[5,10] Few randomized trials have been completed to evaluate the efficacy of one type of treatment (i.e. pharmacotherapy vs. psychotherapy) over another; however, current guidelines state that, since pharmacotherapy has lower effectiveness for PTSD than MH counseling, it is prudent for prescribing clinicians to also refer patients for appropriate counseling.[6,7,10,11]

A literature review helped to identify some factors that are associated with recognition of mental illness and therefore with receipt of treatment for mental illness. First, more severe psychological distress increases the odds that a patient will be diagnosed with a MH disorder.[1,4,12] But little is known about whether this principle applies to PTSD in PC practice. It seems reasonable, however, to suspect that patients with more severe PTSD symptoms would be more likely to receive MH treatment.[1,4,11,12] Secondly, PC physicians are more likely to recognize depressive symptoms (and are more likely to mislabel patients with lone PTSD as having depression).[13] Thus it seems also likely that comorbid depression (with PTSD) might increase the likelihood of receiving MH treatment.[13] Finally, patient disclosure of trauma-associated symptoms to a medical professional seemed to also increase the likelihood of receiving MH treatment.[1,6,7,12–14]

To better understand factors associated with receipt of PTSD treatment by PC patients, a cross-sectional study of patients in PC was conducted assessing PTSD diagnosis and reviewing patients' medical records for MH diagnoses and PTSD treatments. Given the therapeutic overlap (i.e., SSRI pharmacotherapy) in the management of PTSD and depression, it was hypothesized that some patients with PTSD might receive MH treatment despite the fact that they are not recognized as having PTSD. In fact, it was thought that these patients might be mislabeled in the medical record

Avenue, 2nd Floor, Boston, MA 02118, USA.

Ellen C. Meltzer, MD, MSc, Division of Medical Ethics, Weill Cornell Medical College, 435 East 70th Street; Suite 4JM, New York, NY, 10021, USA.

Richard Saitz, MD, MPH, Department of Epidemiology, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118, USA.

Richard Saitz, MD, MPH, Youth Alcohol Prevention Center, Boston University School of Public Health, 580 Harrison Avenue, Boston, MA, USA.

Jeffrey H. Samet, MD, MA, MPH, Department of Community Health Science, Boston University School of Public Health, 801 Massachusetts Avenue, 3rd Floor, Boston, MA 02118, USA.

Jane M. Liebschutz, MD, MPH, Department of Community Health Science, Boston University School of Public Health, 801 Massachusetts Avenue, 3rd Floor, Boston, MA 02118, USA. Phone: +1-617-4147399; Fax: +1-617-4144676; Email: jane.liebschutz@bmc.org Email: jliebs@bu.edu

Khelda Jabbar, MD, Department of Family Medicine, Boston Medical Center, Boston University, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA.

Christine Lloyd-Travaglini, MPH, Data Coordinating Center, Boston University School of Public Health, 801 Massachusetts Avenue, 3rd Floor, Boston, MA 02118, USA.

as having other MH diagnoses (especially depression) and thus receive PTSD treatment fortuitously due to this phenomenon of therapeutic overlap. It was also hypothesized that PTSD symptom severity and trauma-associated symptom disclosure would be associated with receipt of treatment for PTSD.

Methods

Study design

This secondary data analysis examines participants who met diagnostic criteria for current PTSD during a cross-sectional study completed at the PC clinics of an urban, safety-net, academic medical center.[2] This analysis had three key components. First, validated measures established the overall prevalence of interview-diagnosed PTSD, PTSD symptom severity, and depressive symptoms.[15] Second, participants' electronic medical records (EMRs) were reviewed for the presence of four MH diagnoses, SSRI prescriptions, and documentation of visits with MH professional in the prior year. Finally, logistic regression tested the associations between receipt of MH treatment and the hypothesized independent variables. A detailed description of study methods for recruitment and assessment can be found elsewhere; relevant methods are summarized below.[2] The Boston University Medical Center's Institutional Review and the HIPAA Privacy Review Boards approved the study. A Certificate of Confidentiality was obtained from the National Institutes of Health.

Recruitment and enrollment

From February 2003 to September 2004, adult patients awaiting PC appointments were approached and screened for eligibility by trained research assistants. Patients were eligible if they spoke English, were 18–65 years old, and had a scheduled appointment with a PC clinician. Of the 753 eligible patients, 607 (81%) enrolled in the study.[2] This analysis is limited to the 133 patients with a current (past 12 months) diagnosis of PTSD made using the Composite International Diagnostic Interview (CIDI) Version 2.1 PTSD Module.[16,17] Patients who participated in the study provided written informed consent, were compensated $10, and received safety referrals at the end of the interview.

Assessments and data

Interview assessments

Research assistants collected demographic data and administered a series of validated interview questionnaires. They administered CIDI to assess for current PTSD and the Patient Health Questionnaire-9 (PHQ-9) to measure depressive symptoms.[15–17] PTSD symptom severity was ascertained via the PTSD Checklist (PCL-C).[18,19] Participants were asked specifically whether they had ever disclosed that they suffered from trauma-associated symptoms to a medical professional, defined as a PC physician, MH professional/therapist, other physician, nurse, or social worker.

EMR data

All study subjects were patients at an academic medical center which maintains a comprehensive EMR. All outpatient encounters (PC clinical encounters and emergency department visits), inpatient discharge summaries, diagnoses, and prescriptions are documented in the EMR and are available for review. MH services are available at the same academic medical center and visits with

an MH professional are also documented in the EMR; however, restrictions do not allow for the specific contents of these visits (i.e., type of therapy performed) to be viewed. Using standardized data-collection forms, medical students and residents trained in chart abstraction, and supervised by an academic internist, reviewed each participant's EMR, starting from 12 months prior to the date of entry into the study. The EMR was reviewed for MH diagnoses and prior year MH treatments (prescription of an SSRI and/or greater than or equal to one visit with MH professional).[20,21] MH diagnoses included ICD-9-coded PTSD, depression, anxiety, and panic disorder in a patient's problem list and/or in a clinician's typed assessment from any given visit (excluding MH visits). With respect to visits with MH professional, the type of MH professional seen and the type of behavioral therapy received was not available for review. However, MH professional could enter MH diagnosis into a patient's problem list and/or prescribe an SSRI, both of which would appear in the EMR and be available for review for this study.

Main variables

The primary dependent variable was *receipt of MH treatment in the prior 12 months*. MH treatment was defined as receipt of either an SSRI prescription and/or greater than or equal to one visit with a MH professional, as these are treatments that could be effective for PTSD. There were four main independent variables of interest. *MH diagnoses in the EMR* included PTSD, depression, anxiety and/or panic disorder. *More severe PTSD symptoms* was a categorical variable indicating the highest quartile of the distribution of all participants' PCL-C scores.[19] *Comorbid depression* was defined as a PHQ-9 score of nine or greater, based on published scoring cutoffs from the PHQ-9 derived from studies correlating past two week depressive symptoms with a diagnosis of major or other depression.[15] *Disclosure of suffering from trauma-associated symptoms to a medical professional* was a dichotomous variable.[15,18,19] Covariates of interest included the sociodemographic factors age, sex, race (black vs. other), marital status, education level, employment status, and annual income. Insurance status was not included in the model as more than 99% of participants had coverage for the types of utilization studied via federal, state, or private insurance or through an uncompensated care pool ("free care").

Statistical analysis

Descriptive and bivariate analyses were conducted using *t*-tests for continuous data and the Chi-square (χ^2) test for categorical data. To determine factors associated with patients receiving MH treatment, logistic regression was tested for associations between the dependent variable, the main independent variables of interest, and potential confounding factors. The latter were selected if they had statistically significant associations with the outcome in bivariate analyses ($p<0.05$) or, in the cases of more severe PTSD symptoms and comorbid depression, due to prespecified hypotheses.

Results

Characteristics of patients with PTSD

Of the 133 participants (all with current PTSD by diagnostic interview), the mean age was 41 years (SD=11; ***Table 1). Eighty-two (62%) were female and 75 (56%) were black. Almost half had never been married, 40 (30%) had less than 12 years of education, and 83 (62%) were unemployed or on disability. Eighty-five (67%) participants earned less than $20,000 annually. The majority (71%) also had comorbid depression by diagnostic interview.

Table 1

Characteristics of primary care patients with PTSD at an Urban Safety Net Hospital: overall and stratified by receipt of PTSD treatment

	Total (*N*=133)	Any PTSD treatment[4] (*n*=66)	No PTSD treatment (*n*=67)	
Variable	***N* (%)**	***N* (col.%)**	***N* (col.%)**	***p* Value**
Age, years, mean (SD)	41 (11)	42 (11)	39 (11)	0.2
Female gender	82 (62%)	41 (62%)	41 (61%)	0.9
Race				
Black	75 (56%)	30 (45%)	45 (67%)	0.03[5]
White	24 (18%)	17 (26%)	7 (11%)	
Hispanic	13 (10%)	9 (14%)	4 (6%)	
Other	21 (16%)	10 (15%)	11 (16%)	
Marital status				
Married/with live-in partner	24 (18%)	8 (12%)	16 (24%)	0.2
Separated/divorced	38 (29%)	23 (35%)	15 (22%)	
Widowed	8 (6%)	4 (6%)	4 (6%)	
Never married	63 (47%)	31 (47%)	32 (48%)	
Education				
<12 years	40 (30%)	18 (27%)	22 (33%)	0.8
12 years	45 (34%)	23 (35%)	22 (33%)	
>12 years	48 (36%)	25 (38%)	23 (34%)	
Employment status				
Full-time	21 (16%)	5 (8%)	16 (24%)	0.01
Part-time	22 (17%)	8 (12%)	14 (21%)	
Student	7 (5%)	3 (5%)	4 (6%)	
Unemployed	83 (62%)	50 (76%)	33 (49%)	
Income				
<$20,000	85 (67%)	45 (71%)	40 (63%)	0.3
≥$20,000	41 (33%)	18 (29%)	23 (37%)	
Comorbid depression[1]	94 (71%)	45 (68%)	49 (73%)	0.5
Disclosure of trauma-associated symptoms	81 (61%)	52 (79%)	29 (44%)	<0.001
More severe PTSD symptoms	82 (62%)	44 (67%)	38 (57%)	0.2

Table 1
(Continued)

Variable	Total (N=133)	Any PTSD treatment[4] (n=66)	No PTSD treatment (n=67)	
	N (%)	N (col.%)	N (col.%)	p Value
EMR documentation of mental illness				
Any mental illness[2]	83 (62%)	58 (88%)	25 (37%)	<0.001
PTSD	14 (11%)	12 (18%)	2 (3%)	0.004
Anxiety/panic attack	30 (23%)	22 (33%)	8 (12%)	0.003
Depression[3]	66 (50%)	47 (71%)	19 (28%)	<0.001

[1]Depressive symptoms correlate with past 2 weeks major or other depression
[2]Includes PTSD, anxiety, bipolar/manic disorder, panic disorder, major, and other depression
[3]Major and other depression
[4]Includes SSRI and/or visit with MH professional
[5]p Value compares the Black race vs. all other groups combined

EMR documentation of mental illness

Table 1 describes the MH diagnoses in subjects' EMRs. The majority (88%) of participants who received MH treatment had at least one mental illness diagnosis documented in the EMR, most commonly depression (71%). For participants with research interview-diagnosed PTSD alone (N=39) and participants with research interview-diagnosed PTSD and comorbid depression (N=94), frequencies of EMR documentation of PTSD (10% vs. 11%, p=0.9) and depression (44% vs. 52%, p=0.4) were similar.

Receipt of MH treatment

Nearly half of the participants received MH treatment in the prior year: 23 (17%) had both an SSRI prescription and at least one visit with MH professional; 24 (18%) received only an SSRI prescription; and 19 (14%) had no SSRI medication but at least one visit with MH professional (mean number of visits was 4.5 with a range of 1–50; Fig. 1).

Predictors of receipt of MH treatment

In bivariate analyses, age and sex were not significantly different between those who received MH treatment compared to those who did not. Fewer black patients received MH treatment compared to all other racial groups combined, while more Caucasians and Hispanics did receive treatment. A significantly higher proportion of treated patients were unemployed or on disability.

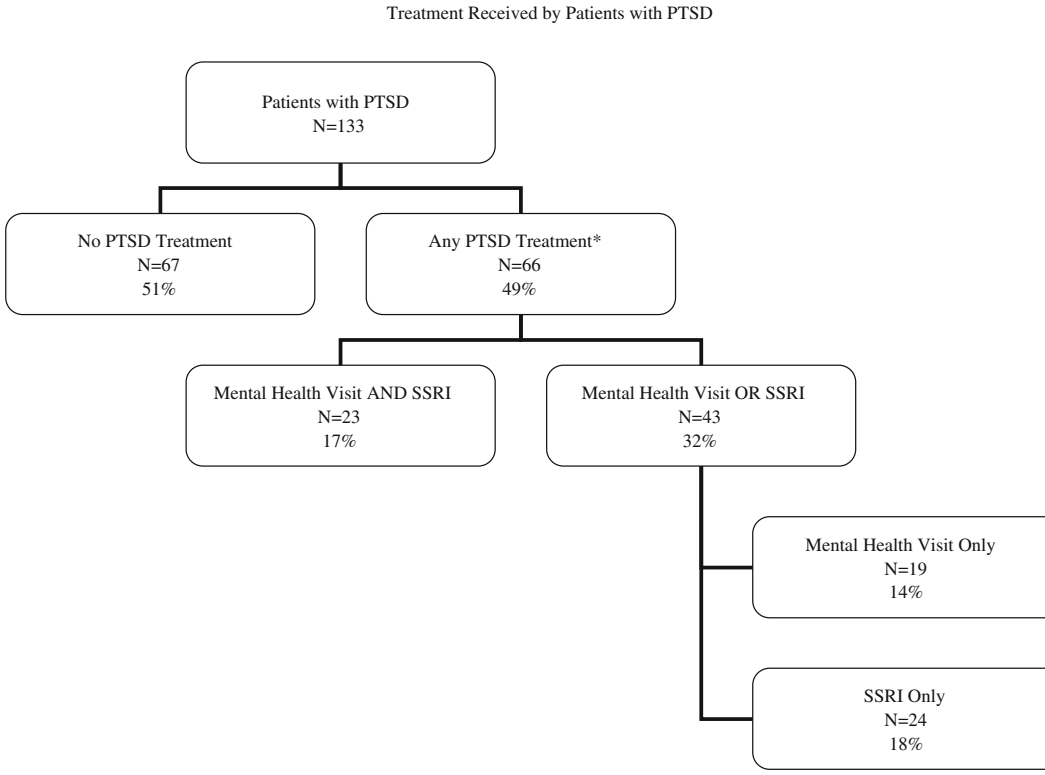


Figure 1

Treatment received by patients with PTSD. * Mean number of visits with MH professional is 4.5; standard deviation 7.6. Range 1–50

Among participants with more severe PTSD symptoms and those with comorbid depression as determined from the research interview, there was no statistically significant difference in the proportion that received MH treatment. Compared to untreated subjects, a higher proportion of treated ones reported that at some point they had disclosed suffering from trauma-associated symptoms to a medical professional. In addition, a significantly higher proportion of treated participants had a diagnosis of PTSD, depression, anxiety, and/or panic disorder in their EMR (Table 1).

In adjusted analyses, the odds of receiving MH treatment were increased 8.2 times (95% CI 3.1–21.5) for participants with an EMR MH diagnosis, even if the patient did not specifically have PTSD documented in the EMR. Disclosure of trauma-associated symptoms to a medical professional increased the adjusted odds of receiving MH treatment by 2.6 (95% CI 1.1–6.4). Being unemployed or on disability was also statistically significant (OR 2.7, 95% CI 1.1–6.7). Although attenuated in adjusted analyses, black patients were less likely to receive MH treatment. Race, as well as other factors hypothesized to be clinically relevant (more severe PTSD symptoms and comorbid depression as determined by the research interview), were not statistically significant in this analysis (Table 2).

Discussion

Among a sample of urban PC patients with PTSD, few patients had this diagnosis listed in their medical record. However, despite not having documented PTSD diagnoses, nearly 50% received MH treatment: either an SSRI and/or a visit with MH professional. In addition to a diagnosis of PTSD, any MH diagnosis (depression, anxiety, panic disorder) in the EMR, disclosure of trauma-associated symptoms to a medical professional, and being unemployed or on disability, were all associated with receipt of MH treatment.

Initially factors that might be associated with receipt of MH treatment were selected from the literature. The first two hypotheses were that having more severe PTSD symptoms and/or having comorbid depression (by research interview) would correlate positively with participants receiving MH treatment. These hypotheses were not supported in either bivariate or adjusted analyses. One potential explanation is that participants who received MH treatment may have improved as a result of the treatment and thus had less severe symptoms at the time of assessment for this cross-sectional study. Another plausible reason for these findings is the cross-sectional design of this study that, by definition, did not allow for prospective observations of participants over time.[14]

Table 2
Factors associated with receiving PTSD treatment[1]

Factor	Odds ratio[2]	95 % CI
EMR MH diagnoses	8.21	3.14.–21.52
Disclosure of trauma–associated symptoms	2.61	1.06–6.43
Unemployed/on disability vs. employed/student	2.68	1.08–6.65
Black race vs. others	0.42	0.18–1.02
Comorbid depression	0.52	0.18–1.49
More severe PTSD symptoms	1.32	0.50–3.51

[1]Adjusted odds ratio via logistic regression modeling including factors statistically significant in bivariate analysis

[2]As compared to patients who did not receive prior year PTSD treatment consisting of SSRI and/or visit with MH professional

Meredith et al.[22] found that lack of time and patient financial burden were the strongest barriers to diagnosis and treatment of PTSD among PC physicians, not doubt of the diagnosis. While prospectively Kessler et al.[14] demonstrated that while some patients with depression and anxiety are not recognized as suffering from mental illness at an initial consultation, over time, they are diagnosed and receive treatment.

The next hypothesis tested was that some patients with PTSD receive MH treatment despite the fact that they are not actually diagnosed as having PTSD, likely due to recognition of a comorbid mental illness or misdiagnosis of PTSD. In this sample, 29% of patients had lone PTSD and 71% had comorbid PTSD and depression by research interview. Interestingly, EMR documentation of PTSD was low among all patients (1 in 10). However, 50% of the sample had depression documentation in the EMR. Patients with both PTSD and depression were more frequently diagnosed as suffering from depression. Patients with lone PTSD (i.e., no depression) were commonly diagnosed incorrectly as suffering from depression. These data are consistent with the findings of Samson et al., who demonstrate that among PC patients with PTSD, it is the symptoms of depression or anxiety that are more likely to be recognized and treated.[13] It may be concluded that, in spite of not being appropriately diagnosed with PTSD, many patients receive some MH treatment simply due to therapeutic overlap in the management of common mental illnesses.[2,3,5,7,8] However, insufficiently treating patients with PTSD may result in partially treated PTSD and its associated comorbidities.[23–25]

While first-line pharmacotherapy for PTSD, depression, and anxiety/panic disorder is a SSRI, consensus statements recommend that patients with PTSD also have specialized MH counseling, with structured cognitive behavioral therapy (CBT) or psychotherapy, as part of a comprehensive treatment plan.[5–10] This contrasts to depression, for which monotherapy with antidepressants is recognized by current guidelines as an effective first-line treatment for patients with mild, moderate, or even severe major depression.[26,27] In clinical trials of patients with PTSD, CBT is superior to SSRI treatment, with 50% of patients achieving remission with lone CBT versus 30% with lone SSRI treatment.[11] Thus consensus guidelines suggest that PC patients with PTSD ought to be referred for psychological treatment.[8,9] In this study, it is likely that most participants who received a visit with MH professional did not receive CBT or other specific psychotherapy, as the mean number of visits (4.5) was fewer than the number of visits typically necessary for CBT in studies of PTSD (9–12 sessions).[28]

Ultimately, there is a significant benefit to focused PTSD treatment, as incomplete treatment leads to suboptimal outcomes. Patients with partially treated PTSD (termed Partial PTSD), although less symptomatic than those who meet criteria for full PTSD, still suffer from clinically meaningful symptoms and are at increased risk for suicide.[3,23–25,29–33] In addition, patients with untreated anxiety disorders, including PTSD, more frequently utilize healthcare and generate substantial direct and indirect costs.[34] The United States Preventive Services Task Force currently does not offer any recommendations on screening for PTSD.[35] Until formal guidelines are available and more physicians are proficient in diagnosing PTSD, the findings of this study suggest that physicians consider inquiry about trauma-associated symptoms for those with anxiety or depressive symptoms, or those in high prevalence populations (e.g., returning military or past substance dependence) to facilitate identification of PTSD and thus referral to appropriate treatment.[36] Prospective trials are needed to establish the evidence base for routine inquiry.

Another finding merits discussion. Being unemployed or on disability was also significantly associated with receipt of MH treatment. Patients with PTSD (and anxiety disorders) as well as more significant impairment (i.e., are on disability) more frequently utilize healthcare.[3,4] Perhaps it is this increase in utilization which, as Kessler et al.[14] suggests, ultimately leads to the accurate diagnosis of PTSD. Once diagnosed, patients with PTSD who are unemployed or disabled, may also be more available to partake in treatment compared to their employed counterparts.[14]

Although the relationship was attenuated when adjusted for other factors, blacks were less likely to be diagnosed with PTSD and less likely to receive MH treatment. Perhaps the fact that cultural differences exist in the experience of psychological trauma can partially explain this phenomenon.[37] Others studies show similar findings and posit that a patient's race may actually affect a physician's awareness of mental illness.[38] Further investigation into this phenomenon is needed in order to develop interventions to reduce this disparity.

Perhaps physicians are hesitant to assign a diagnosis of PTSD for fear that the stigma of a MH diagnosis may exacerbate the symptoms of mental illness or lead to discrimination.[39] However, studies have shown, and physicians should be counseled, that there is a therapeutic benefit for patients in the accurate recognition of mental illness.[12] Ormel et al. examined recognition of mental illness by general practitioners in the Netherlands. In addition to the fact that patients recognized as having mental illness were more likely to be treated, compared to those who were not recognized, recognition on its own had positive effects on patient psychopathology and social functioning.[12] Ormel suggests that the key elements present in the process of mental illness recognition, acknowledgement, re-interpretation, and social support, form the therapeutic basis for these positive results.[12]

This study has several limitations. The study was conducted at an urban, academic safety-net hospital, with the majority of participants unemployed, or on disability, and earning less than $20,000 annually. This may seem to make the results less generalizable to other practice settings. However, there are numerous similar practice settings in US cities. Furthermore, historically studies examining patients with mental illness in other PC settings similarly demonstrate that those with mental illness have higher rates of disability affecting both social and occupational functioning.[40,41] The cross-sectional study design does not allow for patients to be followed over time. This does limit the types of inferences one can make. For example, during the interview, patients were asked whether they have ever disclosed to a medical professional that they suffer from trauma-associated symptoms. For those patients who both answered affirmatively and also received MH treatment, one cannot assume this disclosure was made prior to receiving treatment. Although prior research shows that a known history of trauma is highly correlated with receipt of MH treatment, the temporal relationship in this study is not known and one cannot assume causality.[7] Another aspect of the study design, the EMR review, provided some limitations. There was no data on the length or content of treatment that patients received, nor whether patients were adherent to their treatments. In addition, it is not known whether some patients received MH treatment prior to the year before the study interview. Such information might have captured additional treatments for PTSD, however it is also not known whether patients had a diagnosis of PTSD more than a year prior to the study interview. Thus the added value of such information is unclear (this study examined utilization during the time period that coincides with the diagnostic information available). The EMR did not allow for the capture of information on patients who may have been offered, yet refused, MH treatment. In addition, it is not known if any participants sought MH treatment outside of the study facility. Given that a large percentage were unemployed or on disability and earned less than $20,000 annually, it was speculated that they were less likely to utilize MH services outside of the safety-net hospital. Finally, as there was no access to the content of visits with MH professionals, it was not possible to evaluate the type of therapy that may have been received. However, using validated measures to assess for PTSD at the time of entry into the study allowed for the study of patients with current mental illness and to evaluate the types of treatment they received while they had a documented diagnosis. Most importantly, these data provide new insight into factors that impact whether or not a PC patient with PTSD receives MH treatment, and demonstrate the need for further research examining both patient and physician related barriers to PTSD diagnosis and treatment.

Implications for Behavioral Health

Among urban, safety-net hospital PC patients with current PTSD, half received some sort of MH treatment in the prior year, defined as a SSRI and/or a visit with MH professional. However, rather than a diagnosis of PTSD in their EMRs, many patients were identified as suffering from depression, anxiety, and/or panic disorder. Thus it appears as though treatment was often fortuitously received due to therapeutic overlap in the management of PTSD and other common mental illnesses (specifically depression). Encouraging PC patients to disclose if they suffer from trauma-associated symptoms may improve PC identification and treatment of PTSD. Future research, focused on strategies to reduce patient and physician barriers to disclosing trauma and aimed at improving the diagnosis and treatment of PTSD in PC, is essential to advancing the delivery of MH treatment to those suffering from this disabling condition.

Acknowledgement

The work was supported by a Generalist Physician Faculty Scholar Award, RWJF #045452, from the Robert Wood Johnson Foundation, Princeton, NJ and by a career development award, K23 DA016665, from the National Institute on Drug Abuse, National Institutes of Health (J.M.L.). The authors would like to thank Andrew J. Meltzer, M.D, for providing valuable comments on the manuscript.

Conflict of Interest None

References

1. Taubman-Ben-Ari O, Rabinowitz J, Feldman D, et al. Post-traumatic stress disorder in primary care settings: Prevalence and physicians' detection. *Psychological Medicine* 2001; 31:555–560.
2. Liebschutz J, Saitz R, Brower V, et al. PTSD in urban primary care: High prevalence and low physician recognition. *Journal of General Internal Medicine* 2007; 22(6):719–726.
3. Stein M, McQuaid J, Pedrelli P, et al. Posttraumatic stress disorder in the primary care medical setting. *General Hospital Psychiatry* 2000; 22:261–269.
4. Rodriguez BF, Weisberg RB, Pagano ME, et al. Culpepper L, Keller MB. Mental health treatment received by primary care patients with posttraumatic stress disorder. *Journal of Clinical Psychiatry* 2003; 64(10):1230–1236.
5. Ballenger JC, Davidson JR, Lecrubier Y, et al. A proposed algorithm for improved recognition and treatment of the depression/anxiety spectrum in primary care. *Primary Care Companion to the Journal of Clinical Psychiatry* 2001; 3(2):44–52.
6. Ballenger JC, Davidson JR, Lecrubier Y, et al. Consensus statement update on posttraumatic stress disorder from the international consensus group on depression and anxiety. *Journal of Clinical Psychiatry* 2004; 65 Suppl 1:55–62.
7. Ballenger JC, Davidson JR, Lecrubier Y, et al. Consensus statement on posttraumatic stress disorder from the international consensus group on depression and anxiety. *Journal of Clinical Psychiatry* 2000; 61 Suppl 5:60–66.
8. The expert consensus guideline series. Treatment of Posttraumatic Stress Disorder. The Expert Consensus Panels for PTSD. *Journal of Clinical Psychiatry* 1999; 60 Suppl 16:3–76.
9. Ursano RJ, Bell C, Eth S, et al. Practice guideline for the treatment of patients with acute stress disorder and posttraumatic stress disorder. *American Journal of Psychiatry* 2004; 161(11 Suppl):3–31.
10. American Psychiatric Association. Practice guideline for the treatment of patients with acute stress disorder and posttraumatic stress disorder. In: *American Psychiatric Association practice guidelines for the treatment of psychiatric disorders 2004*. American Psychiatric Publishing, 2011.
11. Kelly J. New approaches help heal combat-related PTSD. *NeuroPsychiatry Reviews* 2006; 7(1):1.
12. Ormel J, Van Den Brink W, Koeter MW, et al. Recognition, management and outcome of psychological disorders in primary care: A naturalistic follow-up study. *Psychological Medicine* 1990; 20(4):909–923.
13. Samson AY, Bensen S, Beck A, et al. Posttraumatic stress disorder in primary care. *Journal of Family Practice* 1999; 48(3):222–227.
14. Kessler D, Bennewith O, Lewis G, et al. Detection of depression and anxiety in primary care: Follow up study. *British Medical Journal* 2002; 325(7371):1016–1017.
15. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. *JAMA: The Journal of the American Medical Association* 1999; 282 (18):1737–1744.
16. World Health Organization. Composite international diagnostic interview (CIDI), Version 2.1. 1997.
17. American Psychiatric Association, Task Force on DSM-IV. *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)* (4th, text revision edition). Washington, DC: American Psychiatric Association, 2000.

18. Lang AJ, Stein MB. An abbreviated PTSD checklist for use as a screening instrument in primary care. *Behaviour Research and Therapy* 2005; 43:585–594.
19. Blanchard EB, Jones-Alexander J, Buckley TC, et al. Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy* 1996; 34(8):669–673.
20. ICD-9-CM. International classification of diseases, 9th revision. Clinical modification (3 rd edition, volumes 1, 2 and 3). Official authorized addendum effective October 1, 1990–HCFA. *Journal of the American Medical Record Association* 1990; 61(8):Suppl 1–Suppl 35.
21. Dendukuri N, McCusker J, Bellavance F, et al. Comparing the validity of different sources of information on emergency department visits: A latent class analysis. *Medical Care* 2005; 43(3):266–275.
22. Meredith LS, Eisenman DP, Green BL, et al. Tobin J. System factors affect the recognition and management of posttraumatic stress disorder by primary care clinicians. *Medical Care* 2009; 47(6):686–694.
23. Mylle J, Maes M. Partial posttraumatic stress disorder revisited. *Journal of Affective Disorders* 2004; 78(1):37–48.
24. Stein MB, Walker JR, Hazen AL, et al. Full and partial posttraumatic stress disorder: Findings from a community survey. *American Journal of Psychiatry* 1997; 154(8):1114–1119.
25. Marshall RD, Olfson M, Hellman F, et al. Comorbidity, impairment, and suicidality in subthreshold PTSD. *American Journal of Psychiatry* 2001; 158(9):1467–1473.
26. Trivedi MH. Using treatment algorithms to bring patients to remission. *Journal of Clinical Psychiatry* 2003; 64 Suppl 2:8–13.
27. American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder. In: *American Psychiatric Association practice guidelines for the treatment of psychiatric disorders 2004* (3 rd edition). American Psychiatric Publishing, 2011.
28. Harvey AG, Bryant RA, Tarrier N. Cognitive behaviour therapy for posttraumatic stress disorder. *Clinical Psychology Review* 2003; 23 (3):501–522.
29. Nicolaidis C, Curry M, McFarland B, et al. Violence, mental health, and physical symptoms in an academic internal medicine practice. *Journal of General Internal Medicine* 2004; 19(8):819–827.
30. Brown PJ, Stout RL, Gannon-Rowley J. Substance use disorder-PTSD comorbidity. Patients' perceptions of symptom interplay and treatment issues. *Journal of Substance Abuse Treatment* 1998; 15(5):445–448.
31. Stewart SH, Pihl RO, Conrod PJ, et al. Functional associations among trauma, PTSD, and substance-related disorders. *Addictive Behaviors* 1998; 23(6):797–812.
32. Brown PJ, Wolfe J. Substance abuse and post-traumatic stress disorder comorbidity. *Drug and Alcohol Dependence* 1994; 35(1):51–59.
33. Grinage BD. Diagnosis and management of post-traumatic stress disorder. *American Family Physician* 2003; 68(12):2401–2408.
34. Greenberg PE, Sisitsky T, Kessler RC, et al. The economic burden of anxiety disorders in the 1990s. *Journal of Clinical Psychiatry* 1999; 60(7):427–435.
35. Guide to Clinical Preventive Services, 2008. *Recommendations of the U.S. Preventive services task force. AHRQ publication no. 08–05122*, September 2008.
36. Doyle B, Foa EB, Keane TM, et al. Posttraumatic stress disorder: Changing the perspective of primary care. *Medical Crossfire* 2001; 3 (4):33–43.
37. Watters E. Suffering differently. *New York Times Magazine*, April 12, 2007.
38. Borowsky SJ, Rubenstein LV, Meredith LS, et al. Who is at risk of nondetection of mental health problems in primary care? *Journal of General Internal Medicine* 2000; 15(6):381–388.
39. Corrigan PW. How clinical diagnosis might exacerbate the stigma of mental illness. *Social Work* 2007; 52(1):31–39.
40. Olfson M, Fireman B, Weissman MM, et al. Mental disorders and disability among patients in a primary care group practice. *American Journal of Psychiatry* 1997; 154(12):1734–1740.
41. Kroenke K, Spitzer RL, Williams JB, et al. Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine* 2007; 146(5):317–325.

Pain Medicine

Pain Medicine 2012; 13: 1436–1443
Wiley Periodicals, Inc.

OPIOIDS, SUBSTANCE ABUSE & ADDICTIONS SECTION

Original Research Article

Aberrant Drug-Related Behaviors: Unsystematic Documentation Does Not Identify Prescription Drug Use Disorder

Ellen C. Meltzer, MD, MSc,*† Dennis Rybin, MS,¶ Lidia Z. Meshesha, BA,* Richard Saitz, MD, MPH,*‡ Jeffrey H. Samet, MD, MA, MPH,*§ Sonia L. Rubens, MA,*†† and Jane M. Liebschutz, MD, MPH*§**

*Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center and Boston University School of Medicine, Boston, Massachusetts;

†Division of Medical Ethics, Weill Medical College of Cornell University, New York, New York;

Departments of ‡Epidemiology and

§Community Health Sciences;

¶Data Coordinating Center and

**Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, Massachusetts;

††Clinical Child Psychology Program, University of Kansas, Lawrence, Kansas, USA

Reprint requests to: Jane M. Liebschutz, MD, MPH, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA. Tel: 617-414-6910; Fax: 617-414-4676; E-mail: jane.liebschutz@bmc.org.

Disclosures: The study was funded by K23 DA016665 from the National Institute of Drug Abuse. Dr. Jane Liebschutz had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. This work was previously presented at the *Society of General Internal Medicine 32nd Annual Meeting* in Miami, Florida on May 15, 2009, where it was a finalist for the Mack Lipkin Sr. Associate Member Awards.

Conflict of Interest: The authors all report no conflicts of interest. R. Saitz is a consultant for Medical Directions Inc. and BMJ Publishing Group.

All work was performed at Boston University Medical Center.

Abstract

Objective. No evidence-based methods exist to identify prescription drug use disorder (PDUD) in primary care (PC) patients prescribed controlled substances. Aberrant drug-related behaviors (ADRBs) are suggested as a proxy. Our objective was to determine whether ADRBs documented in electronic medical records (EMRs) of patients prescribed opioids and benzodiazepines could serve as a proxy for identifying PDUD.

Design. A cross-sectional study of PC patients at an urban, academic medical center.

Subjects. Two hundred sixty-four English-speaking patients (ages 18–60) with chronic pain (≥3 months), receiving ≥1 opioid analgesic or benzodiazepine prescription in the past year, were recruited during outpatient PC visits.

Outcome Measures. Composite International Diagnostic Interview defined *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) diagnoses of past year PDUD and no disorder. EMRs were reviewed for 15 prespecified ADRBs (e.g., early refill, stolen medications) in the year before and after study entry. Fisher's exact test compared

frequencies of each ADRB between participants with and without PDUD.

Results. Sixty-one participants (23%) met DSM-IV PDUD criteria and 203 (77%) had no disorder; 85% had one or more ADRB documented. Few differences in frequencies of individual behaviors were noted between groups, with only "appearing intoxicated or high" documented more frequently among participants with PDUD (N = 10, 16%) vs no disorder (N = 8, 4%), $P = 0.002$. The only common ADRB, "emergency visit for pain," did not discriminate between those with and without the disorder (82% PDUD vs 78% no disorder, $P = 0.6$).

Conclusions. EMR documentation of ADRBs is common among PC patients prescribed opioids or benzodiazepines, but unsystematic clinician documentation does not identify PDUDs. Evidence-based approaches are needed.

Key Words. Prescription Drug Use Disorder; Diagnosis; Aberrant Drug-Related Behaviors; Primary Care; Chronic Pain

Introduction

Chronic pain is a common presenting problem in the primary care (PC) setting; approximately 22% of PC patients report chronic pain [1]. Over the past two decades, opioid prescribing for chronic noncancer pain has dramatically increased, accompanied by rising rates of opioid misuse, unintentional overdose, and legal prosecution of physicians [2–6]. Benzodiazepine prescribing has similarly become a topic for debate, not only due to risks for medication abuse and diversion but also because of the association of benzodiazepines to opioid-related deaths [7]. Given that PC physicians (PCPs) provide the vast majority of psychoactive substances nationally [8], they must take great care in the decision to prescribe these medications and monitor for side effects to avoid such negative outcomes. In addition, PCPs must assess for risk of prescription drug use disorder (PDUD), defined as abuse of or dependence on prescription opioids and prescription benzodiazepines and/or any illicit substance and/or alcohol dependence while receiving opioid and/or benzodiazepine pharmacotherapy [9,10]. One suggested method of monitoring patients prescribed opioids and benzodiazepines is assessing for aberrant drug-related behaviors (ADRBs) [9].

Experts in pain medicine define ADRBs as behaviors suggesting out of control use of medications, one hallmark of addiction. Examples of such behaviors include insisting on a medication by name, claiming non-narcotic medications "do not work," buying medications off the street, making frequent emergency visits for pain medications, asking for an early refill, and spending extensive time discussing medications [11]. ADRBs may also include taking a medication in a manner not prescribed—e.g., taking someone else's medication, unsanctioned dose escalations, multiple prescription locations, or different providers writing prescriptions [11]. Finally, in addition to behaviors that are suggestive of a patient taking medication in an out of control manner, some behaviors suggest medication diversion, the transfer of legally obtained drugs into illegal channels—where patients exchange or sell their prescription drugs to someone else [4,12]. It is worth noting that when ADRBs do occur, there can be multiple possible explanations for the behavior and hence consideration of a differential diagnosis of their etiology is appropriate, with PDUD being one of several plausible explanations [13].

Currently, experts recommend monitoring patients prescribed chronic opioids for ADRBs based on the assumption that ADRBs indicate medication misuse, diversion, or addiction [14–19]. However, the evidence supporting this assumption is very limited: studies using ADRBs have relied on physician recall and/or documentation of ADRBs in a patient's chart as a proxy for PDUD [20–23]. In order to establish an evidence base for the use of ADRBs as a proxy for PDUD in patients with chronic pain prescribed opioids and benzodiazepines, we compared chart documentation of ADRBs with a systematically obtained patient diagnosis of PDUD.

Methods

Study Design

This was a cross-sectional study of PC patients with chronic pain who were recruited from the PC clinics of an urban, academic, safety-net medical center [24,25]. Safety-net hospitals in the United States care for poor and vulnerable populations who may be uninsured or underinsured and include disproportionate numbers of underrepresented minorities [26]. We collected data from participants by an in-person interview with a trained member of the research team and by a subsequent electronic medical record (EMR) review for abstraction of prescription opioid and benzodiazepine data to assess entry criteria, as well as to identify any documentation of 15 prespecified behaviors considered by experts to be concerning for addictive disease [10,11]. Further specific details of study methods have been published previously [24,25].

Setting

Research interviewers recruited patients waiting for scheduled PC visits. Interviewers were physicians, master-level professionals, college graduates, and college students who underwent 60+ hours of interview training. All participants were approached in the waiting rooms of the hospital-based PC practice [25].

Recruitment and Enrollment

Of the 833 (40.0%) patients eligible for the study based on explicit criteria, i.e., were 18–60 years of age, spoke English, endorsed pain of at least 3-month duration, reported use of any analgesic medication, including over

the counter or prescription, in the prior month, and had a scheduled PC appointment, 589 patients (70.7%) agreed to participate. Informed consent was obtained from eligible patients, and participants were compensated $10 for the one time interview. The Boston University Medical Center Institutional Review Board approved the study.

All EMR entries from 12 months prior to study entry were reviewed looking for documentation of any opioid or benzodiazepine prescription; a comprehensive list of all included medications has been previously published [24,25]. Standardized chart abstraction forms were used. EMRs were comprehensive including notes from all clinic visits, all emergency department records, all inpatient discharge summaries, phone calls, and an institutionalprescription database. This data analysis was based on the 264 (44.8%) study participants who were prescribed an opioid pain reliever and/or a benzodiazepine in the 12 months prior to their interview.

Measures and Key Variables

Independent Variable: PDUD

PDUD was defined as per DSM-IV diagnostic criteria for current (past year) abuse of or dependence on prescription opioids and prescription benzodiazepines [10]. Also included in this definition was any illicit substance and/or alcohol dependence while receiving opioid and/or benzodiazepine pharmacotherapy. During the interview portion, participants were assessed using the Composite International Diagnostic Interview (CIDI) v.2.1 module on Drug Disorders to diagnose drug use disorders [27]. Criteria for drug abuse included social, physical, or legal consequences from use. The criteria for drug dependence included compulsive use, health consequences, and physical dependence (i.e., tolerance or withdrawal). Physical dependence alone did not suffice to meet the diagnosis. The CIDI short form was used to measure current alcohol dependence; current alcohol abuse and past alcohol use disorders were not measured in order to reduce respondent time burden [27]. Thus, participants with PDUD may have a lone prescription opioid and/or benzodiazepine disorder, a lone substance use disorder (SUD) (an illicit drug disorder and/or past year alcohol dependence) while receiving prescription opioids and/or benzodiazepines, or a comorbid prescription opioid and/or benzodiazepine disorder along with another SUD (i.e., comorbid illicit drug disorder and/or past year alcohol dependence) [27].

While it may be questioned whether a lone SUD while receiving prescription opioids and/or benzodiazepines truly constitutes PDUD, this is a definition that has been used clinically and also suggested by experts in pain and addiction [10,28,29]. However, this was taken into account and a sensitivity analysis was conducted, in which those with a lone other SUD while prescribed an opioid and/or benzodiazepine were analyzed as a unique group. As this analysis yielded similar results, for simplicity the data presented are from a two-group analysis comparing those with PDUD as defined above to those with no disorder. Nicotine dependence was not included in the variable SUD.

Dependent Variable: ADRBs

Two years of EMR data were reviewed for each participant in order to assess ADRBs, records from 12 months prior to and poststudy entry, looking for documentation of 12 prespecified behaviors. These behaviors were recognized in the published literature as signs of potential addiction or diversion in patients prescribed controlled substances [11]. While our list of behaviors was based on the work of Portenoy and meant to represent the concepts he first put forward, our nomenclature was slightly modified. Standardized chart abstraction forms were used.

Statistical Analysis

Using Fisher's exact test, frequencies of each ADRB and cumulative numbers of ADRBs were compared between participants with PDUD and no disorder.

Results

The demographic characteristics of the 264 participants, stratified by the presence or absence of PDUD, are presented in Table 1. Twenty-three percent (61/264) of study participants met criteria for current PDUD. We found few demographic differences between the two groups (i.e., age, gender, race, and education). Characteristics included mean age of mid-40s, majority African American, and over two-thirds with 12 or more years of education. At least 50% of each group, with and without PDUD, were receiving disability payments and over 95% of each group had severely disabling pain [30]. However, participants with PDUD were more likely to have the following characteristics: posttraumatic stress disorder (PTSD), depression, smoker, past time in jail, and a family history of substance abuse (Table 2).

Among patients receiving prescriptions for opioids in the prior 12 months, 15% received the equivalent of 20 tablets of 5 mg oxycodone in ≤2 fills, 12.6% received 21–60 tablets in ≤3 fills, 22.7% received 61–150 tablets in ≤3 fills, and 49.6% received >150 tablets or >3 fills of any amount (e.g., four prescriptions of 20 tablets each). The majority of those in the last category received >6 fills. Of those receiving benzodiazepines in the prior 12 months in one or more prescriptions (N = 66), 17% received less than 30 pills, 15% received 30–100 pills, 15% received 101–200 pills, and 41% received >200 pills. Eight patients (12%) received benzodiazepine prescriptions from outside psychiatrists, for which specific pill counts and fills were unavailable.

Of the 15 behaviors assessed, only "appearing intoxicated or high" was documented more frequently in the EMRs of participants with PDUD (N = 10, 16%) vs no disorder (N = 8, 4%), $P = 0.002$. However, the frequency was low and this behavior was present in the charts of those with no disorder (Table 2). Indeed most participants in both groups

Table 1 Demographic characteristics of PC participants with chronic pain and analgesic medication use stratified by prescription drug use disorder (PDUD) (N = 264)

Variable		PDUD N = 61 (23%)	No Disorder N = 203 (77%)	*P* Value
Mean age (year)		45.9	47.3	0.24
Female gender		49%	58%	0.24
Race	African American	54%	60%	0.26
	Hispanic	13%	9%	
	White	28%	20%	
	Other	5%	11%	
Education	>12 years	66%	73%	0.26
On disability		57%	51%	0.47
PTSD		49%	33%	0.02
Depression		56%	33%	0.002
Current smoker		70%	45%	<0.001
Ever in jail		71%	33%	<0.001
Family substance abuse		74%	48%	<0.001

PC = primary care; PTSD = posttraumatic stress disorder.

had at least one ADRB (85%, *P* = 1.0). With respect to the remaining 14 behaviors, the frequency of any one documented behavior was not significantly different among patients with PDUD and with no disorder. An urgent visit for pain was the most common aberrant behavior, but frequency did not differ between the two groups (PDUD vs no PDUD—82% vs 78%, *P* = 0.6). Early refill was present in about 15% of both groups of patients. The remaining behaviors were present in 15% or less of participants. With respect to cumulative behaviors, patients with PDUD do not have significantly more documented ADRBs in their EMR (Figure 1). Rather, having one or two behaviors in the chart was common; over 80% of the patients with and without PDUD had at least one behavior in their chart and almost 40% in both groups had two ADRBs.

Discussion

By using patients' EMRs and research interviews, this study investigated whether ADRBs documented in patients' charts reflected a diagnosis of PDUD. If documented ADRBs were indeed viable indicators of PDUD, then the EMR abstractions would have revealed significant differences in patients' charts; patients with a diagnosis of PDUD would have had more ADRBs documented in their medical charts than those with no diagnosis. The data reveal no differences; ADRBs as noted in routine PC practice do not identify patients with PDUDs. Despite the thoughtful clinical insight that went into the development of these recommended "clinical pearls," empirically they do not yield the desired outcome, identification of patients that should be of great concern to the clinician. However, the negative results of this study nevertheless do provide insight about physician recording in the EMR and suggest that relying on nonsystematic documentation to identify addiction or diversion is not useful.

Table 2 Frequency of aberrant drug-related behaviors stratified by prescription drug use disorder (PDUD) (N = 264)

Behavior	PDUD N = 61 N (%)	No Disorder N = 203 N (%)	*P* Value
Appears intoxicated/high	10 (16%)	8 (4%)	0.002
Use someone else's medication	3 (5%)	8 (4%)	0.8
Bought medication off the street	3 (5%)	3 (2%)	0.1
Extensive time discussing medication	3 (5%)	9 (4%)	1.0
Involvement in an accident	3 (5%)	29 (14%)	0.07
Number of prescription locations	3 (5%)	4 (2%)	0.2
Insists non-narcotics do not work	1 (2%)	8 (4%)	0.7
Try to get scripts from other MDs	1 (2%)	3 (2%)	1.0
Urgent visit for pain	50 (82%)	159 (78%)	0.6
Early refill	9 (15%)	29 (14%)	1.0
Urgent visit for narcotic pain medications	9 (15%)	14 (7%)	0.06
Insists on medication by name	7 (12%)	22 (11%)	0.8
Lost medication	5 (8%)	11 (5%)	0.5
Increase dose without authorization	4 (7%)	12 (6%)	0.6
Reports stolen medication	3 (5%)	3 (2%)	0.1

MD = medical doctor.

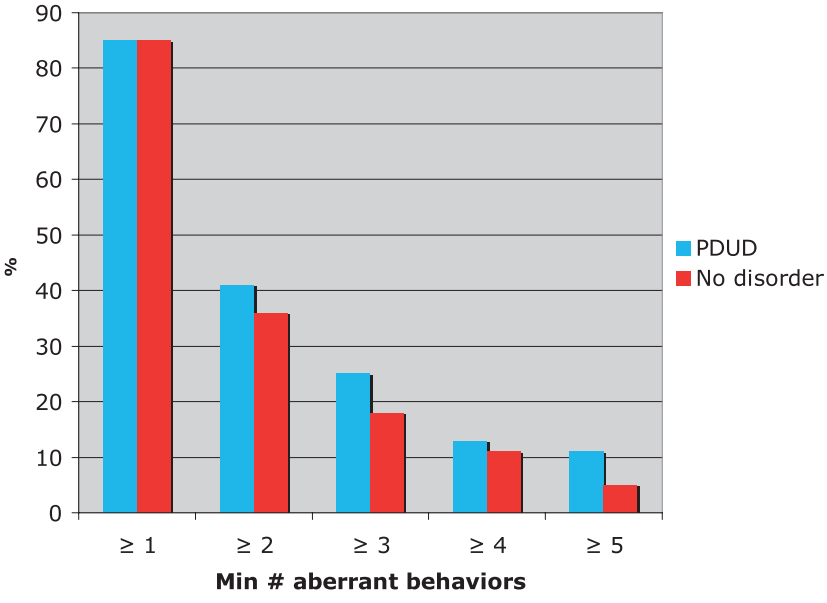


Figure 1 Cumulative aberrant drug-related behaviors among participants with prescription drug use disorder (PDUD) and no disorder. No results were statistically significant.

Clinicians prescribing opioids for patients with chronic pain are aware that some degree of ADRBs may occur during the course of treatment. Accordingly, the provider should be aware that there is a differential diagnosis for ADRBs, of which PDUD is only one potential explanation [31]. Included in this differential is pain that is not adequately controlled, leading patients to increase their dose of medication without authorization and the subsequent need for an early refill [31]. For example, randomized controlled trials have demonstrated that opioid analgesia is not consistently effective; thus, patients with uncontrolled pain may take additional medication in an effort to find relief [32,33]. In addition, some patients with adequately controlled pain may hoard medications, fearful of a time when the pain may suddenly worsen [11,34]. This too can lead to early refills and even use of multiple prescription locations. Finally, a pain flare can send a patient to the emergency department in an effort to find relief. Thus, patients without PDUD can also plausibly have ADRBs documented in their charts. This study found that this is in fact the case, and these events do occur at a comparable frequency as ADRBs in patients with and without PDUD. The presence of these behaviors in those *without* PDUDs has been referred to as a pseudoaddiction, in which the aberrant behaviors will resolve once the pain is adequately treated [13]. It is clear that the patients in this study did not have adequately controlled pain, as over 95% of both groups of patients had severely disabling pain. Perhaps if the pain were better treated, then a true difference in documented ADRBs would exist between the groups. Alternatively, as suggested by Portenoy and McCarberg and Passik, perhaps many of the behaviors considered to be concerning are really more benign and not truly associated with addiction or diversion [11,31].

One documented behavior warrants a deeper investigation. From the 15 ADRBs examined, only appearing intoxicated or high was documented more frequently in the EMRs of patients with PDUD. Any time a clinician suspects a patient of being intoxicated or high, further evaluation is warranted. Recently published guidelines offer clinicians a framework for safely initiating and continuing to prescribe opioids [9]. It is suggested that at every clinical visit, the patient is assessed for any risks associated with the opioids as well as any benefits received from the medication. Prescribing ought to continue when the benefits of a medication outweigh its potential risks, thus ensuring an ethically equitable distribution of benefits and burdens [35]. If a patient seems to be intoxicated or high, concern arises that the potential risks to the patient, as well as society, outweigh any benefits. In this situation, titration off of the opioid or benzodiazepine combined with the provision of adjunctive therapy for pain and/or anxiety, as well as referral to substance abuse treatment and/or psychiatry, is likely most appropriate [36].

The other 14 ADRBs were not significantly different between the groups. As this study examined documented behaviors, the negative results suggest that relying on unsystematic physician documentation is not useful for identifying patients with PDUD. Rather, in order to determine if ADRBs can be useful to monitor patients for PDUDs, a reasonable next step would be to assess implementation of standardized clinical assessment tools into clinical practice.

Although it may appear that ADRBs are not useful in detecting PDUD, the nonsystematic collection of ADRBs may be the problem rather than the ADRBs themselves [37]. Validated tools have been studied to predict misuse of substances while being prescribed opioid analgesia for chronic noncancer pain [9]. For example, the Screener and Opioid Assessment for Patients with Pain™ is a tool to help risk-stratify patients at the

initiation of opioid therapy [20]. The results provide an idea of how closely a patient ought to be monitored during treatment [20]. In addition, the Current Opioid Misuse Measure (COMM) provides a means for ongoing risk estimation via standardized assessment of aberrant behaviors [38]. The present authors have studied the COMM and caution that clinical assessment tools may perform differently in distinct patient populations, as the sensitivity and specificity of some tools may vary with the prevalence of disease [24]. As this study demonstrates, ADRBs as recorded in a patient's EMR are not associated with PDUD. Thus, caution should always be utilized when monitoring any patient for the development of PDUD, even when standardized assessment tools are in place. As Butler et al. stress, the purpose of these tools is not to serve as the basis for punitive measures [38]. Rather, they are to help clinicians have a means for consistent patient assessment and to limit the need to rely on haphazard documentation.

Finally, the results demonstrate that patients with PDUD were more likely to suffer from PTSD and/or depression, to be current smokers, to have ever been in jail, and to have a family history of substance abuse. The present authors have previously explored these clinical risk factors for PDUD [25]. In addition, as previously reported, over 30% of patients without PDUD also suffered from comorbid mental illness, smoked, had been in jail, and had a family history of substance abuse [25]. Clearly, there is a significant burden of comorbid mental illness in patients with chronic pain, offering further support of the need for an interdisciplinary approach to the management of chronic pain [36].

This study has limitations. All participants with PDUD may not have been properly identified through inadequacy of the study instruments or through inaccurate reporting on the part of study participants. This would lead to misclassification bias, in which patients with PDUD are inappropriately labeled as having no disorder, serving to bias the results toward the null hypothesis [39]. In addition, tolerance to and withdrawal from opioids and benzodiazepines can be a naturally occurring phenomenon but are nonetheless included in the diagnostic interview for PDUD [10]. This would serve to bias results away from the null hypothesis. Finally, the CIDI does not assess for drug diversion, a specific risk related to prescription opioids and benzodiazepines [5,12,40]. Even with its acknowledged limitations, the CIDI has been used widely and is considered to be a well-validated diagnostic instrument [27]. The sample size was small and included only 61 patients with PDUD. Even if a larger sample size revealed statistically different findings, individual ADRBs would likely have little predictive value in diagnosing PDUD, as the majority of patients without a substance disorder also demonstrated each of the behaviors. Two years worth of chart entries were reviewed for each study participant, thus providing data to examine cumulative numbers of behaviors exhibited, which were not significant. However, when we looked at a combined number of behaviors, we did not stratify them and look at a hierarchy of more severe behaviors. Overall, the results contribute to the recognition of the potential limitations of documented ADRBs and suggest a need to utilize a standardized approach to patient assessment as a potentially more informative way to gain clinical insight from ADRBs.

Conclusion

Among PC patients with chronic pain receiving prescription opioids and/or benzodiazepines, having at least one ADRB documented in the EMR is almost universal. In addition, frequencies of most individual behaviors are similar among those with and without PDUD. Reliance on nonsystematic documentation of ADRBs to identify PDUD in PC patients with chronic pain may not be useful. Based on these findings, physicians and researchers should be cautious before using documented ADRBs in the EMR as a proxy for PDUD. Prospective studies, in which ADRBs are systematically assessed, are needed. Only then can we determine the true significance of these behaviors.

References

1 Gureje O, Von Korff M, Simon GE, Gater R. Persistent pain and well-being: A World Health Organization study in primary care. [erratum appears in JAMA 1998 Oct 7;280(13):1142]. JAMA 1998;280(2):147–51.

2 Marks RM, Sachar EJ. Undertreatment of medical inpatients with narcotic analgesics. Ann Intern Med 1978;78:173–81.

3 Cicero TJ, Inciardi JA, Munoz A. Trends in abuse of oxycontin and other opioid analgesics in the United States: 2002–2004. J Pain 2005;6(10):662–72.

4 Bell J. The global diversion of pharmaceutical drugs: Opiate treatment and the diversion of pharmaceutical opiates: A clinician's perspective. Addiction 2010;105(9):1531–7.

5 Centers for Disease Control and Prevention (CDC). Emergency department visits involving nonmedical use of selected prescription drugs—United States, 2004–2008. MMWR Morb Mortal Wkly Rep 2010;59(23):705–9.

6 Reidenberg MM, Willis O. Prosecution of physicians for prescribing opioids to patients. Clin Pharmacol Ther 2007;81(6):903–6.

7 Webster LR, Cochella S, Dasgupta N, et al. An analysis of the root causes for opioid-related overdose deaths in the United States. Pain Med 2011;12(suppl 2):S26–35.

8 Smith DE. Prescribing practices and the prescription drug epidemic: Physician intervention strategies. J Psychoactive Drugs 2012;44(1):68–71.

9 Chou R, Fanciullo GJ, Fine PG, et al. Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. J Pain 2009;10(2):113–30.

10 American Psychiatric Association, Task Force on DSM-IV. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), 4th edition. Washington, DC: American Psychiatric Association; 2000:xxxvii, 943.

11 Portenoy RK. Opioid therapy for chronic nonmalignant pain: A review of the critical issues. J Pain Symptom Manage 1996;11(4):203–17.

12 Inciardi JA, Surratt HL, Cicero TJ, et al. The "black box" of prescription drug diversion. J Addict Dis 2009;28(4):332–47.

13 Weissman DE, Haddox JD. Opioid pseudoaddiction—An iatrogenic syndrome. Pain 1989;36(3):363–6.

14 Dunbar SA, Katz NP. Chronic opioid therapy for nonmalignant pain in patients with a history of substance abuse: Report of 20 cases. J Pain Symptom Manage 1996;11(3):163–71.

15 Passik SD, Kirsh KL, McDonald MV, et al. A pilot survey of aberrant drug-taking attitudes and behaviors in samples of cancer and AIDS patients. J Pain Symptom Manage 2000;19(4):274–86.

16 Michna E, Ross EL, Hynes WL, et al. Predicting aberrant drug behavior in patients treated for chronic pain: Importance of abuse history. J Pain Symptom Manage 2004;28(3):250–8.

17 Ives TJ, Chelminski PR, Hammett-Stabler CA, et al. Predictors of opioid misuse in patients with chronic pain: A prospective cohort study. BMC Health Serv Res 2006;6:46.

18 Fleming MF, Balousek SL, Klessig CL, Mundt MP, Brown DD. Substance use disorders in a primary care sample receiving daily opioid therapy. J Pain 2007;8(7):573–82.

19 Kirsh KL, Whitcomb LA, Donaghy K, Passik SD. Abuse and addiction issues in medically ill patients with pain: Attempts at clarification of terms and empirical study. Clin J Pain 2002;18(4 suppl):S52–60.

20 Wasan AD, Butler SF, Budman SH, et al. Does report of craving opioid medication predict aberrant drug behavior among chronic pain patients? Clin J Pain 2009;25(3):193–8.

21 Butler SF, Budman SH, Fernandez K, Jamison RN. Validation of a screener and opioid assessment measure for patients with chronic pain. Pain 2004;112(1–2):65–75.

22 Weaver MF, Bond DS, Arnold BL, Waterhouse E, Towne A. Aberrant drug-taking behaviors and headache: Patient versus physician report. Am J Health Behav 2006;30(5):475–82.

23 Compton P, Darakjian J, Miotto K. Screening for addiction in patients with chronic pain and "problematic" substance use: Evaluation of a pilot assessment tool. J Pain Symptom Manage 1998;16(6):355–63.

24 Meltzer EC, Rybin D, Saitz R, et al. Identifying prescription opioid use disorder in primary care: Diagnostic characteristics of the current opioid misuse measure (COMM). Pain 2011;152(2):397–402.

25 Liebschutz JM, Saitz R, Weiss RD, et al. Clinical factors associated with prescription drug use disorder in urban primary care patients with chronic pain. J Pain 2010;11(11):1047–55.

26 Baxter RJ, Mechanic RE. The status of local health care safety nets. Health Aff (Millwood) 1997;16(4):7–23.

27 Kessler RC, Andrews G, Mroczek D, Ustun B, Wittchen HU. The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF). Int J Methods Psychiatr Res 1998;7(4):171–85.

28 Zacny J, Bigelow G, Compton P, et al. College on problems of drug dependence taskforce on prescription opioid non-medical use and abuse: Position statement. Drug Alcohol Depend 2003;69(3):215–32.

29 Katz NP, Adams EH, Benneyan JC, et al. Foundations of opioid risk management. Clin J Pain 2007;23(2):103–18.

30 Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992;50(2):133–49.

31 McCarberg B, Passik SD. Expert Guide to Pain Management. Philadelphia, PA: American College of Physicians; 2005.

32 Martell BA, O'Connor PG, Kerns RD, et al. Systematic review: Opioid treatment for chronic back pain: Prevalence, efficacy, and association with addiction. Ann Intern Med 2007;146(2):116–27.

33 Ossipov MH, Lai J, King T, Vanderah TW, Porreca F. Underlying mechanisms of pronociceptive consequences of prolonged morphine exposure. Biopolymers 2005;80(2–3):319–24.

34 Portenoy RK. Opioid therapy for chronic nonmalignant pain: Clinician's perspective. J Law Med Ethics 1996;24(4):296–309.

35 Jonsen AR, Siegler M, Winslade WJ. Clinical Ethics: A Practical Approach to Ethical Decisions in Clinical

Medicine, 7th edition. New York: McGraw-Hill Medical; 2010.

36 Liebschutz JM, Alford DP. Safe opioid prescribing: A long way to go. J Gen Intern Med 2011;26(9): 951–2.

37 Krebs EE, Ramsey DC, Miloshoff JM, Bair MJ. Primary care monitoring of long-term opioid therapy among veterans with chronic pain. Pain Med 2011;12(5): 740–6.

38 Butler SF, Budman SH, Fernandez KC, et al. Development and validation of the current opioid misuse measure. Pain 2007;130(1–2):144–56.

39 Fletcher RH, Fletcher SW. Clinical Epidemiology the Essentials, 4th edition. Baltimore, MD: Lippincott Williams & Wilkins; 2005.

40 McCabe SE, Teter CJ, Boyd CJ. Medical use, illicit use, and diversion of abusable prescription drugs. J Am Coll Health 2006;54(5):269–78.

ann. behav. med. (2012) 43:62–73
DOI 10.1007/s12160-011-9319-4

ORIGINAL ARTICLE

A Culturally Adapted Telecommunication System to Improve Physical Activity, Diet Quality, and Medication Adherence Among Hypertensive African–Americans: A Randomized Controlled Trial

Jeffrey P. Migneault, Ph.D. · Julien J. Dedier, M.D., M.P.H. · Julie A. Wright, Ph.D. · Timothy Heeren, Ph.D. · Marci Kramish Campbell, Ph.D. · Donald E. Morisky, Sc.D. · Peter Rudd, M.D. · Robert H. Friedman, M.D.

Published online: 14 January 2012

Abstract

Background Hypertension is more prevalent and clinically severe among African–Americans than whites. Several health behaviors influence blood pressure (BP) control, but effective, accessible, culturally sensitive interventions that target multiple behaviors are lacking.

Purpose We evaluated a culturally adapted, automated telephone system to help hypertensive, urban African–American adults improve their adherence to their antihypertensive medication regimen and to evidence-based guidelines for dietary behavior and physical activity.

Methods We randomized 337 hypertensive primary care patients to an 8-month automated, multi-behavior intervention or to an education-only control. Medication adherence, diet, physical activity, and BP were assessed at baseline and every 4 months for 1 year. Data were analyzed using longitudinal modeling.

Results The intervention was associated with improvements in a measure of overall diet quality (+3.5 points, $p<0.03$) and in energy expenditure (+80 kcal/day, $p<0.03$). A decrease in systolic BP between groups was not statistically significant (−2.3 mmHg, $p=0.25$).

Conclusions Given their convenience, scalability, and ability to deliver tailored messages, automated telecommunications systems can promote self-management of diet and energy balance in urban African–Americans.

Study registered at ClinicalTrials.gov with ID: NCT00207194

J. P. Migneault · J. J. Dedier (✉) · T. Heeren · R. H. Friedman
Medical Information Systems Unit, General Internal Medicine,
Boston University School of Medicine,
Boston, MA, USA
e-mail: jdedier@bu.edu

J. A. Wright
Department of Exercise and Health Science,
University of Massachusetts,
Boston, MA, USA

M. K. Campbell
Department of Nutrition, Gillings School of Global Public Health,
University of North Carolina,
Chapel Hill, NC, USA

D. E. Morisky
Department of Community Health Sciences,
UCLA School of Public Health,
Los Angeles, CA, USA

P. Rudd
Department of Medicine, Stanford University School of Medicine,
Stanford, CA, USA

Keywords Behavioral intervention · Cultural adaptation · Hypertension

Introduction

In the USA, the prevalence of hypertension is significantly higher among African–Americans compared to Whites. Although the majority of US Blacks who have hypertension receive medical care and antihypertensive medication, only 58% of them achieve adequate blood pressure control [1].

Multiple well-designed studies have shown that lifestyle interventions delivered in-person can improve adherence to antihypertensive medication regimens, to regular physical activity, to a healthful diet, and to weight control recommendations—all of which can reduce elevated blood pressure [2–4]. Nonetheless, these programs are not suitable

for use on a public health scale for reasons including their high cost of personnel, facilities, and materials and the burden on participants to attend multiple counseling sessions in the face of other life priorities. Minority groups of low socioeconomic status (SES) may encounter additional barriers including logistical obstacles (e.g., unreliable transportation), cost considerations, and cultural incongruence of intervention content or delivery [5, 6].

The current public health challenge is to design and implement health programs that can be delivered easily to the millions of patients with hypertension in a practical, low-cost manner. Tailoring these programs to the needs, barriers, and ethno-cultural characteristics of specific populations can help maximize their impact [7].

Studies from our group and others have shown that computer-based, telephone intervention systems can be efficacious in changing patients' health-related behaviors, including medication adherence for hypertensive patients [8], physical activity [9, 10], and dietary behaviors [11]. However, most of these have targeted single rather than multiple behaviors, and few have been culturally adapted. The purpose of this study was to evaluate an automated, culturally adapted, multi-behavior intervention that can be used on a public health scale to promote hypertension self-management in urban African–Americans of low socioeconomic status. The primary study hypothesis was that the intervention group would experience greater improvements in these behaviors than the control group. Our secondary hypothesis was that mean blood pressure (BP) would decrease more in the intervention than control group.

Methods

We conducted a two-armed randomized controlled trial (intervention vs. usual care control) of a novel automated telephone counseling system designed to promote three health-related behaviors that affect blood pressure control: (a) adhering to the antihypertensive medication regimen, (b) following the dietary approaches to stop hypertension (DASH) diet [2], and (c) engaging in regular physical activity [12].

Participants were drawn from the adult primary care practices of a large, urban safety-net hospital and from four affiliated community health centers. Inclusion criteria were as follows: (a) African–American by self-report; (b) a diagnosis of hypertension on the active problem list of the patient's medical chart; (c) a current prescription for ≥1 antihypertensive medications; (d) ≥1 primary care office visits in the previous 2 months; (e) two elevated clinic BP readings in the previous 6 months (systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg among non-diabetic patients, and ≥130/80 among diabetic patients); and (f) age ≥35 years. Patients were ineligible if they were pregnant, or had a medical condition that precluded moderate exercise or for which the DASH diet was contraindicated. Patients were also excluded if they scored 7 out of 7 on a modified version of the 8-item Morisky Medication Adherence Scale, indicating high baseline medication adherence [13], had a terminal illness, psychiatric or cognitive disorder, or no access to a telephone.

Potentially eligible patients were identified by searching the electronic medical records of the participating clinical sites for patients who met the inclusion criteria (a–f) above. Primary care providers were then asked to review lists of their potentially eligible patients and remove those with any exclusion. The remaining patients were sent a letter on behalf of their primary care provider that described the study and were asked to return a postage-paid postcard if they did not want to be contacted by study staff. Those who did not opt-out were contacted by phone, the study was explained to them, and they were screened for eligibility. All participants provided informed consent. Based on power analyses and projected attrition, we sought to randomize 360 patients expecting 300 to complete the 8-month study assessment thus providing sufficient power to analyze the three primary behavioral outcomes.

Intervention: Telephone-Linked Care for Hypertension in African–Americans

The intervention was a totally automated, computer-based, interactive telephone counseling system called Telephone-Linked-Care, designed to monitor, educate, and counsel African–American adults with hypertension and to provide summary data regularly to the patient's primary care provider [8–10]. The intervention incorporated principles of social cognitive theory [14], the transtheoretical model of behavioral change [15], and motivational interviewing [16], and was tailored to the user's values. Content was also adapted to cultural characteristics of culturally African–American adults (i.e., not Caribbean–American, Black–Hispanic, etc.). This group-level cultural adaptation was guided by a conceptual framework targeting "surface" and "deep" structural domains, and by ethnic mapping [17] in focus groups of potential users. S*urface structure* adaptation involves matching intervention materials and messages to observable, "superficial" characteristics of a target population including familiar or preferred people, places, language, music, food, locations, and clothing. In this study, surface adaptation included using pre-recorded speech from African–American voice professionals to deliver the automated telephone counseling messages. *Deep structure* can be viewed as the underlying rationale for surface structure preferences. It denotes the cultural values, social, historical, and psychological forces that influence how a given health behavior is viewed by the target population. For example, inviting an African–American adult

to join a neighborhood walking group may have greater deep structure appeal than suggesting he/she walk alone due to the cultural priority of community well-being observed in African–American subgroups [18].

In ethnic mapping, we presented focus group members with foods, activities, and health recommendations and asked to rate them on a continuum from "Mostly a black thing" (e.g., dancing) to "Mostly a white thing" (e.g., skiing, yoga) and we included material in automated telephone scripts, accordingly. These scripts were then pre-tested by potential users of the system and their feedback was incorporated. New script elements emerged from focus group material related to topics such as dance as an effective method to combine fun and healthy exercise, and using low-cost community resources, home-based exercise, or mall-walking to achieve a safe place to exercise. Another new element addressed concerns over "sweating out" one's hairdo as a barrier to exercise:

> "Some hairstyles can get messed when you sweat or swim in a pool. It's nice to look good, but it's also important to be healthy. Consider a low-maintenance style that will look good on you and hold up well with an active lifestyle, like a short natural cut or braids. Ask your hairstylist's opinion the next time you visit. And remember—regular exercise can help you look good too."

The automated telephone intervention delivered one call per week for 32 weeks. The first three calls introduced the three targeted behaviors and their role in BP control, and described the system. Subsequent calls were arranged as modules on medication adherence, physical activity, and diet, and were delivered in the order chosen by the participant. Each call consisted of (a) an introduction, (b) a section for reporting health information collected on study-issued home measurement devices (pedometers, sphygmomanometers, weight scales), and (c) theory-based interactive education and counseling on the targeted behavior.

The physical activity module consisted of 12 calls to increase levels of moderate-or-greater intensity physical activity and was adapted from our previously evaluated automated telephone systems to promote physical activity [9, 10]. The diet module consisted of nine calls—one overview call and two calls for each of four topics: fruits and vegetables, fiber, sodium, and fat. The content of these calls was designed to promote the DASH diet [2]. The medication adherence module consisted of eight calls and was adapted from a successful prototype from our group [8]. Study staff monitored participant use of the system and contacted those who did not call to assist or re-engage them with the system. In addition, participants and their primary care providers received printouts of the participant's progress in each health behavior. These were sent at the beginning and end of each of the three behavioral modules and were designed to reinforce the intervention.

Randomization Groups

Patients who screened eligible by telephone attended an in-home study visit for the purpose of health education, baseline assessments, and randomization. At this visit, participants received a 75-page resource manual that described hypertension, listed dietary recommendations, heart healthy food recipes, and local resources for exercise, and provided information to support antihypertensive medication adherence. They received a 20-min education session based on the content of this manual, and were given a pedometer and a digital weight scale (Healthometer, model # HDR900KD01). Participants completed baseline study assessments and were then randomized to the automated, telephone-linked behavioral intervention plus standard primary medical care, or to the control condition of standard primary medical care alone. Randomization was accomplished using a random number generator to assign subjects to one of the two groups. Neither participants nor research assistants knew the group assignment until after baseline assessments were complete. After group assignment was revealed, intervention group members received a digital home BP monitor (Omron, model # HEM711AC), were trained to use it, and were trained to use the automated telephone system.

Assessments

Assessments were completed by trained research assistants at the participant's home except for a handful done in the study office for logistical reasons. The three primary study outcomes were change in (a) diet quality, (b) leisure time physical activity of moderate-or-greater intensity, and (c) adherence to the antihypertensive medication regimen from baseline to the end of the 8-month intervention period. The secondary outcome was change in BP. All outcomes were assessed at baseline (randomization), 4, 8 (end-of-intervention), and 12 months. Other assessments included weight, height, psychosocial characteristics, and users' perceptions of the automated telephone system (see Table 1 for a list of demographic, medical, and social characteristics assessed, and for information about how they were measured and categorized).

Physical activity outcomes were assessed with the interviewer-administered 7-day physical activity recall [19]. The main physical activity outcomes were (a) minutes per week of moderate-or-greater intensity physical activity, (b) total energy expenditure per day (kcal/day) [20], and (c) whether the participant met the joint Centers of Disease Control and Prevention—American College of Sports Medicine recommendation of at least 150 min of moderate-or-greater physical activity per week [12]. In order to validate self-reported physical activity, 48 participants were randomly selected to wear an accelerometer (CSA Actigraph,

Table 1 Characteristics of the sample at baseline by randomization group

Variable	Group	
	Automated telephone intervention (n=169)	Control (n=168)
Gender (% female)	65.7	75.0
Age, years, mean (SD)	56.3 (10.6)	56.8 (11.4)
Education, years, mean (SD)	12.0 (2.5)	12.4 (2.8)
Body mass index, kg/m^2, mean (SD)	34.4 (8.6)	34.3 (8.4)
Married or living together (%)	35.7	34.7
Employed full- or part-time (%)	38.9	40.4
Median household income	$10–$20 K/year	$10–$20 K/year
Income perception (%)		
Comfortable with enough for "extras"	19.2	20.3
Enough for bills but not "extras"	19.2	20.3
Cut back in order to pay bills	23.7	24.2
Not enough to pay all bills	37.8	35.3
Insurance (%)		
Medicare/medicaid	53.8	49.4
Free care[a]	19.5	25.3
Other[b]	24.9	24.7
Self-pay	1.8	0.6
Medications, total number[c], mean (SD)	5.1 (3.0)	5.0 (3.3)
Medications, number for blood pressure, mean (SD)	1.9 (0.8)	1.9 (0.8)
Diabetes (%)	40.7	35.8
Stroke (%)	11.2*	4.2*
Co-morbidity index[d], mean (SD)	5.8 (3.8)	6.2 (4.7)
Health literacy score[e], mean (SD)	51.7 (20.5)[g]	54.3 (18.4)[g]
Primary outcomes		
Diet index score, mean (SD)	53.9 (17.6)	55.8 (17.0)
Moderate-or-greater intensity leisure time physical activity, min/week, mean (SD)	162.4 (169.0)*	126.3 (144.3)*
>150 min per week (%)	38.5*	26.2*
Energy expenditure, kcals/day	3234.7 (860.7)	3188.5 (820.3)
Medication Adherence Score[f], mean (SD)	4.93 (1.6)	4.77 (1.4)
Secondary outcomes		
Systolic blood pressure, mean (SD)	130.6 (19.8)	131.8 (18.6)
Diastolic blood pressure, mean (SD)	80.9 (12.5)	80.3 (11.8)

*p<0.05

[a] Health Safety Net (Free Care) funding program for underinsured Massachusetts state residents

[b] Employee-sponsored, self-insured, and other products

[c] Total number of self-reported prescription medications

[d] Oliver Co-morbidity Index [33]

[e] REALM questionnaire [34]

[f] From the 7-item version of the 8-item Morisky Medication Adherence Scale [13]

[g] Scores correspond to a 7th–8th grade reading level

model # W7164LS) for 3 days before the baseline and 8-month assessments.

Diet outcomes were assessed with the picture-sort Food Frequency Questionnaire [21]. Five diet outcomes were derived and were analyzed as continuous variables and as categorical variables denoting whether or not the participant reported consuming recommended levels. The outcomes and their categorical cut-offs are as follows [22]: (a) percent of total calories derived from fat (≤30%); (b) percent of total calories derived from saturated fat (≤7%); (c) grams of fiber (≥25 g/day); (d) servings of fruit and vegetables (F&V ≥5 servings); and (e) milligrams of sodium (NA ≤2,400 mg/day).

A composite diet quality score was constructed that combined the five diet outcomes using an approach similar to that used in previous research [11]. Each diet outcome variable was converted into a score from 0 to 100, with higher scores indicating healthier dietary intake. Participants received a score of 100 if they met or exceeded the dietary recommendations described above for total fat, saturated fat, and fiber. However, for fruit and vegetables and sodium intake, a score of 100 points was based on stricter DASH diet recommendations [2]: ≥10 fruit and vegetable servings/day and ≤1,500 mg NA/day. As there are no published criteria to anchor the unhealthy end of the five component dietary outcome scales listed above, we assigned a score of zero to values ≥95th percentile for fat and sodium intake in baseline data, and ≤5th percentile for fiber and F&V. These values were 50.5% of calories from total fat, 16.9% of calories from saturated fat, 9.2 g/day of fiber, 1.5 servings/day of F&V, and 7,663 mg NA/day. This approach captured most of the variation in the sample and reduced the effect of outliers. Next, scores between 0 and 100 were set using a linear transformation (e.g., a value halfway between the criteria for 0 and 100 would receive a score of 50). The overall Diet Quality Score was set as the

mean of these five component scores and, therefore, also had a possible range of 0–100.

Medication adherence was assessed using a 7-item version of the 8-item Morisky Medication Adherence Scale [13]. Scores could range from 0 to 7, with higher scores indicating better adherence.

BP was measured with the Welch Allyn Spot Vital Signs Monitor (Model # 4020-00) by trained research assistants during each home visit using a research protocol similar to other studies [23]. These BP monitors were calibrated at the beginning of the study and annually thereafter. Participants were instructed to sit quietly for 5 min before each reading. At the first home visit a reading was taken in each arm, and the arm with the highest blood pressure was chosen for all subsequent readings. At each visit, three readings were taken in the designated arm and the mean of the last two was used for analysis. Systolic (SBP) and diastolic blood pressure (DBP) were analyzed as continuous outcome variables in separate statistical models. Each was also analyzed as the dichotomous variable ("high" or "controlled") according to national guidelines [24]. Specifically, among non-diabetic participants SBP values ≥140 mmHg and DBP values ≥90 mmHg were considered "high". The corresponding BP cut-off for diabetic participants was 130/80 mmHg. We examined another dichotomous BP outcome that considered BP "controlled" when both the SBP and DBP values were below the recommended cut-off, and "high" when one or both of these values exceeded its cut-off.

Data Analysis

The distributions of the primary outcome variables were examined and adjustments were made to normalize the data when appropriate. For diet outcomes, we eliminated as unreliable all Food Frequency Questionnaire results where the total caloric intake appeared unrealistically low (<800 cal/day) or high (>6,000 cal/day). Consequently, 15 (1.3%) were eliminated for low values and 30 (2.6%) for high values. Results from the remaining sample across the five dietary behaviors were normally distributed except for a few high outliers. These remaining outliers were capped at three standard deviations from the mean, which affected only 32 (0.6%) of the 5,705 total observations.

The distribution of minutes of moderate-or-greater physical activity had a large positive skew; therefore, log-transformed values were used in analyses. Non-transformed means are presented for ease of interpretation. At baseline, the distribution of total energy expenditure (kcal/day) and medication adherence scores were approximately normal.

Preliminary analyses found a higher rate of drop out in the intervention group than among controls, and our primary outcome analyses used an intent-to-treat approach to minimize this potential bias [25]. Missing data were imputed using the last value carried forward. For cases where data were available at time points before and after the missing value, the mean of these two values was used. Results of analyses that considered only participants who were assessed at the 8-month, end-of-intervention time point (i.e., the available-cases analyses) were similar to results from the intent-to-treat analyses.

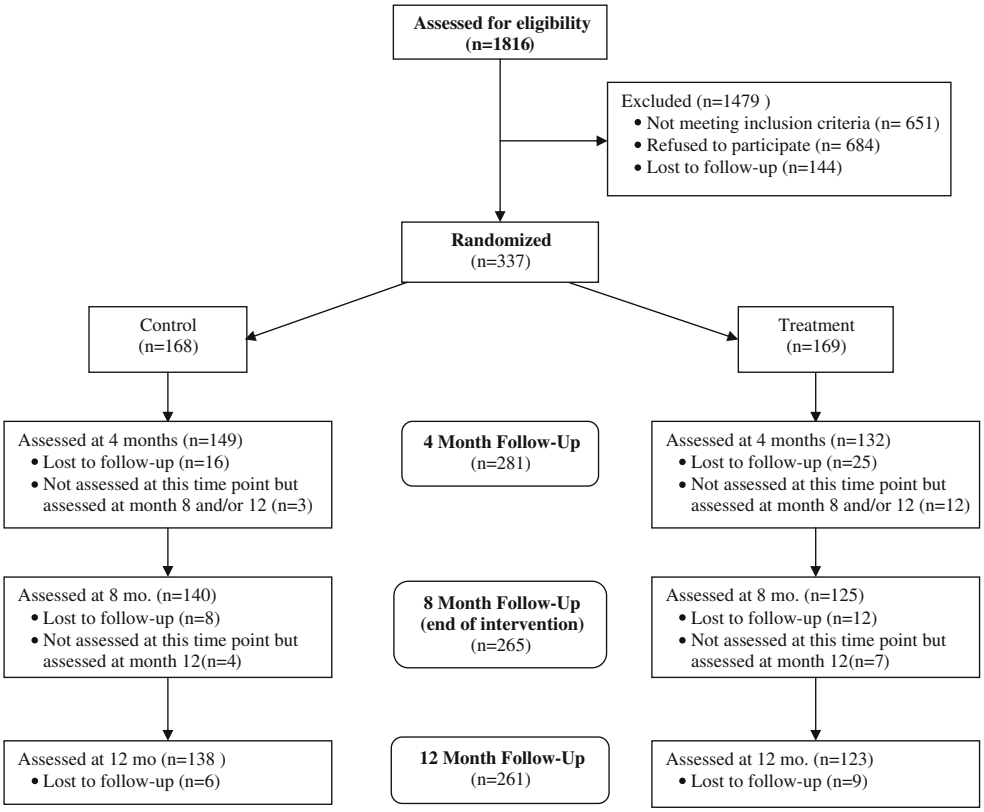
The two study groups were compared on demographic and outcome variables at baseline using t tests for continuous characteristics and chi-square tests for categorical characteristics. Regression models for longitudinal data were used to analyze the effect of randomization group on the change in primary outcomes from baseline to each of the three follow-up assessment time points (i.e., 4, 8, and 12 months). As we hypothesized that intervention effects would be greatest at the end of the intervention, our primary focus was on changes in outcome through the 8-month assessment. Variables on which the groups differed significantly at baseline (e.g., stroke history) or that were likely to independently affect the outcome were included as covariates in each analysis. These included gender, age, perceived financial well-being, education, diabetic status, and stroke history for all analyses (see Table 1). Other covariates were included for particular outcomes as appropriate (e.g., season for physical activity and diet outcomes). Random effects linear regression models were used for continuous outcome measures. Generalized estimating equations logistic regression models were used for categorical outcomes. SAS version 9 was used for all analysis. Dose–response analyses were conducted for participants randomized to the treatment group by calculating the Pearson correlation between the number of completed automated telephone calls and outcomes at 8 months. Analyses were repeated separately for the subset of patients with type 2 diabetes.

Results

Study Participants

Participant recruitment and follow-up occurred from October 2003 to July 2007. Figure 1 displays the flow of potentially eligible participants through the study. Of the 1,816 individuals who were contacted by phone and invited to participate in the study, 684 (35.8%) refused to participate, 651 (37.7%) did not meet eligibility criteria, and 144 (7.9%) were lost to follow-up before the baseline assessment. Of those not meeting eligibility requirements, 161 (24.7%) were excluded because of a perfect medication adherence score on the 7-item Morisky Medication Adherence Scale at the eligibility screening telephone call. The remaining 337 (18.6%) were confirmed to be eligible and were randomized. Table 1 displays sample characteristics at baseline. Of those

Fig. 1 Consort diagram



randomized, 70% were female, 35% had a live-in partner and 37% reported a diagnosis of type 2 diabetes. Of the 99% who reported health insurance, 79% had Medicare/Medicaid or state-subsidized coverage. The sample's SES status was low: 84% had an annual income <$30,000, 27% worked full time, and the average duration of education was 12.2 years. Mean medication adherence was low in both groups, but almost 20% of participants had an adherence score of six or greater at baseline indicating higher adherence [13]. Intervention group participants reported more moderate-or-greater physical activity per week than controls (162.4 min. vs. 126.3 min., $p=0.04$), and a greater percentage of the intervention group met national moderate-or-greater physical activity recommendations (38.5% vs. 26.2%, $p<0.02$). In addition, more intervention group participants than controls reported a history of stroke (11.2% vs. 4.2%, $p<0.02$).

Participant Attrition

By the 8-month assessment, 61 (18.1%) of the 337 randomized participants had become lost to follow-up resulting in a sample size of 276. Sixty-seven percent ($n=41$) of these had dropped out by the 4-month visit. The dropout rate was higher in the intervention group than in the control group (21.9% vs. 14.3%, $p=0.07$), and all of these differential dropouts occurred by the 4-month assessment. We compared dropouts to completers on the characteristics displayed in Table 1 and found that participants who dropped out by the 8-month assessment were more likely to be male (44.3% vs. 26.4% of completers, $p=0.006$). Those who met recommendations for physical activity at baseline dropped out at the same rate across groups, but those who did not meet the recommendation dropped out at twice the rate from the intervention group as from the control group (25.0% vs. 12.1%, $p=0.01$). Lastly, we examined if those who dropped out from the intervention group were different from control group dropouts on other factors. The only significant difference was the control group dropouts were more likely to be diabetic than intervention group dropouts (50% vs. 22.2%, $p=0.03$).

Use of the Automated Telephone System

Overall use of the automated telephone system was modest: 26 (15.4%) did not call the system at all, and 8 (4.5%) connected with it but did not complete a call. The mean number of completed automated telephone calls was 9.0 of the 32 planned calls, and for those who completed at least one call the mean number of completed calls was 11.4. Eight percent of the intervention group completed at least 80% of the 32 planned calls. Participants with diabetes completed 24% more calls than non-diabetic participants (10.8 vs. 8.2, $p=0.08$). Use of the three modules was nearly identical. Among intervention group members, 75 (44.4%) completed at least one diet module call. For medication adherence and physical activity the numbers were 75 (44.4%) and 77 (45.6%), respectively. Participants completed a mean

of 5.03 of 9 diet calls (56%), 5.85 of 12 physical activity calls (49%), and 5.05 of 8 medication adherence calls (63%). Their choice of which module to do first was equally distributed across modules.

We examined whether any of the following 11 patient characteristics predicted patterns of system use: age, gender, income, perceived financial well-being, education, health literacy, number of medications, presence of diabetes or stroke, total number of comorbidities, employment status, presence of a live-in partner, and number of household occupants. We found that none of these variables significantly predicted whether or not an individual used the system. Among intervention group participants who completed at least one automated telephone call, age (r=0.22, p=0.009) and greater perceived financial well-being (r=0.17, p=0.05) correlated positively with the number of calls made. In a multiple linear regression model containing all 11 patient characteristics, only perceived financial well-being (p=0.02) was an independent predictor of number of completed automated telephone calls.

Multivariable Analysis of Outcome Measures

Table 2 presents the change in adjusted means and proportions from baseline to the 8-month (end of intervention) assessment for the primary and secondary outcome variables by study group. No significant intervention effects were observed at the 4- and 12-month assessments. Adjusted group means for continuous outcome variables across all four assessment time points are graphically presented in Fig. 2.

Dietary Outcomes

At the 8-month assessment, there was a significant improvement of 3.5 points (or 0.25 SD, p<0.03) in the overall diet quality score in the intervention group vs. controls. Among the five diet component subgroups, fiber intake improved significantly in the intervention group vs. controls (relative improvement of 2.29 g/day, p<0.02). Favorable changes were also observed for other diet components, but none of these reached statistical significance.

Physical Activity

Actigraph readings from the subsample at baseline and at 8 months correlated with self-reported leisure time moderate-or-greater physical activity (r=0.33, p<0.05), thus providing objective validation of the 7-Day Physical Activity Recall questionnaire data. There was no significant difference between groups in leisure time moderate-or-greater physical activity or in the percentage meeting the recommended

Table 2 Change in the primary and secondary outcomes from baseline to the end of the intervention period (month 8) within and between groups

Variables	Within-group change scores from baseline to the end of intervention		Difference in change scores between groups
	Automated telephone intervention (n=169)	Control (n=168)	
Primary outcomes[d]			
Overall diet quality score	2.8	−0.74	3.54*
MOD+PA[a] (min/week)	−3.44	2.77	−6.21
>150 min/week MOD+PA (%)	−2	5	7
Total energy expenditure (kcal/day)	43.8	−36.2	80.0*
Medication Adherence Score[b]	0.45	0.26	0.19
Secondary outcomes[d]			
Systolic blood pressure (mmHg)	−2.06	0.25	−2.31
Achieved control[c] (%)	0.7	1.9	−1.2
Diastolic blood pressure (mmHg)	−1.28	−0.1	−1.18
Achieved control[c] (%)	6.4	5.1	1.3
Achieved control of both systolic and diastolic blood pressure (%)	5.8	5.0	0.8

*p<0.05

[a] Mod+PA—moderate-or-greater intensity leisure time physical activity

[b] From the 7-item version of the 8-item Morisky Medication Adherence Scale [13]

[c] Systolic blood pressure control defined as <140 mmHg (<130 mmHg if diabetic). Diastolic blood pressure control defined as <90 mmHg (<80 mmHg if diabetic)

[d] Means and proportions are adjusted for covariates used in the specific regression model (see text)

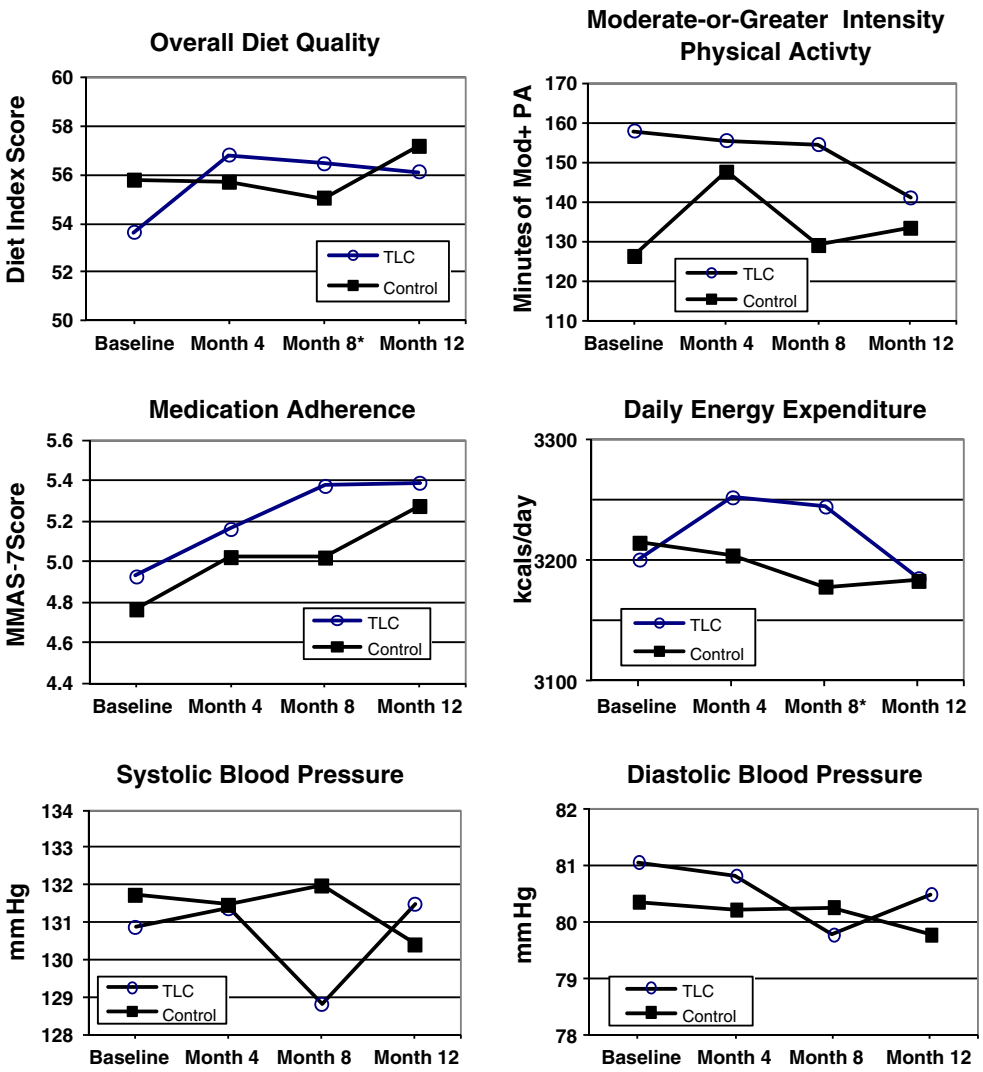


Fig. 2 Primary and secondary outcomes across all time points. *Asterisks* difference between groups in change in outcome from baseline to 8 months significant at $p<0.05$. *TLC* telephone-linked-care, the name of the automated telephone intervention

150 min/week, but there was a significant increase in energy expenditure of 80 kcal/day in the intervention group vs. controls at the 8-month assessment ($p=0.02$). This finding was primarily due to a relative increase in work-related activity in the intervention group.

Medication Adherence

Although the treatment group's adjusted 7-item Morisky Medication Adherence Scale scores improved by 0.19 points relative to controls, this change was not statistically significant ($p=0.25$) in the intent-to-treat analysis. When this analysis was done on available cases, the relative change was larger, at 0.35 points ($p=0.08$).

Blood Pressure

When analyzed as continuous variables, the largest decrease in both SBP and DBP in the intervention group vs. the control group occurred at the 8-month assessment (-2.3 mmHg, and -1.2 mmHg, respectively); however, neither difference was statistically significant. Similarly, there was no statistical difference between groups in the proportion who achieved control of SBP, DBP, or both.

Diabetics

Though there were no pre-specified hypotheses for diabetic participants, we examined intervention effects in this subgroup both because the automated telephone intervention contained material tailored specifically for diabetics with hypertension, and because BP control is especially important in this population. In the intervention group, participants who reported type 2 diabetes used the automated telephone system more and had a lower attrition rate than those without diabetes. Among all diabetics, the intervention was associated with greater absolute improvements in most primary and secondary outcomes relative to controls than for non-

diabetics, but the differences were not statistically significant given the small subgroup size (n=127). For example, in the full sample there was no appreciable effect on minutes of moderate-or-greater physical activity, but among diabetics there was an increase of 37 min/week in the intervention group vs. controls at 8 months (p<0.18). Similarly, among diabetics mean SBP decreased by 6.21 mmHg and DBP 1.51 mmHg in the intervention group vs. controls (p=0.1 and p=0.43, respectively).

Dose Response

There was a dose–response relationship between the number of completed diet calls and improvement in the overall diet quality score (r=0.31, p=0.008). There were no significant dose–response relationships for the other outcomes.

Acceptability of the Automated Telephone System

At the 8-month assessment, participants who completed at least one intervention call were surveyed on their experience with the automated system [26]. Responses to key summary questions were on a scale from 1 to 5 with 5 being the most favorable rating. The mean score for the question "How satisfied were you with the automated telephone system?" was 3.7. For "How helpful did you find the system?" it was 4.1, and for "How helpful did you find the information the system provided?" it was 4.2. The most frequent reasons reported for missing or skipping automated calls were participant related such as not being home when the system called (77%), being too busy (56%), or having major life disruptions (36%). By contrast, system-related reasons for missed calls were reported less often. Examples include "The system wasn't right for me" (16%), and participants having problems accessing the system (14%). Several baseline participant characteristics correlated with users' opinions of the automated telephone system. Those with lower educational attainment rated it as more enjoyable, helpful, useful, friendly, personal, and informative (r=0.19–0.29, p=0.002–0.02). After controlling for educational attainment, the system was rated as more flexible by users with lower vs. higher perceived financial well-being (r=0.19, p=0.04).

Discussion

The intervention was associated with statistically significant improvements in the primary outcomes of overall diet quality and energy expenditure, and with modest improvements in BP that were not statistically significant (Fig. 2). For all outcomes, positive effects did not persist 4 months after the intervention had ended. There was a dose–response effect on diet quality in the intervention group. Diabetic subjects in the intervention group had a significantly lower dropout rate than non-diabetics. They also tended towards greater use of the intervention and greater improvements in most outcomes than non-diabetics though these differences were not statistically significant.

In prior studies from our group, single automated telephone interventions targeting physical activity, diet, or medication adherence alone were each found effective [8, 9, 11]. Although the current intervention combined content from these successful single-behavior prototypes, the observed effects were smaller overall. This difference may relate to differences in the study sample (e.g., proactive recruitment in this study versus volunteer-based in our prior studies) and study protocols (greater participant burden and attrition due to the need to assess multiple behaviors vs. a single behavior). It may also reflect fundamental differences in adherence caused by the effort and complexity of making more than one lifestyle change [27]. Nevertheless, a recent meta-analysis found that computer-tailored interventions targeting up to three health behaviors can be as effective as those targeting only one [28].

With respect to multi-behavior interventions, a further consideration is the relative impact of delivering each module sequentially vs. simultaneously. The scant literature on this issue is mixed. In one study comparing simultaneous and sequential interventions among 289 blacks with hypertension, there was evidence that the simultaneous approaches were superior for sodium reduction and smoking cessation [29]. In contrast, findings from another trial among 315 female smokers supported a sequential over a simultaneous approach to multiple behavior changes including diet, exercise, and smoking cessation [30]. In our study we delivered each module sequentially, but required all participants to check and report their home blood pressure readings throughout follow-up. A subsample also chose to measure and report their weight regularly. Such longitudinal self-monitoring constituted a "simultaneous" intervention that spanned the three behavioral modules and may have helped participants form connections between the content of each module.

In this study, maximal intervention effects were seen at 8 months for several outcomes despite the fact that each behavioral module lasted only 2–3 months. This suggests that subsequent modules may have reinforced positive changes initiated during earlier modules, even though each module targeted a different behavior. This phenomenon is akin to the "co-variation" of effects described in other multi-behavior interventions where improvements in a targeted health behavior are associated with improvements in other targeted or even non-targeted behaviors [31]. Indeed, the first three automated calls were designed not only to introduce the three behavioral modules, but also to emphasize the complementary role of each behavior in blood pressure control. Making these explicit connections

between the three behavioral modules early may have enabled mutual reinforcement of behavioral messages later on. More research is needed to elucidate how messages targeting one behavior can be used to reinforce messages targeting other behaviors in multi-behavior change interventions whether modules are delivered sequentially or simultaneously.

What is clear from the literature and from our results is that positive behavioral changes decay toward baseline levels once the intervention is withdrawn [32]. A more intensive intervention or booster sessions following an intervention can slow this decay [32]. An advantage of automated counseling systems such as the one we report here is that they can be easily modified to deliver a more intensive intervention (i.e., to increase content and contact frequency), and to help maintain positive behavior change by providing booster contacts after the primary intervention ends. In retrospect, providing intermittent reinforcement of positive behavior changes throughout the intervention may have maintained or increased earlier improvements and been more effective.

The low overall rates of system use in this study likely limited intervention effects. Factors such as lower income and educational attainment and the presence of social–environmental obstacles have been shown to predict lower subject participation in research studies [5, 6]. There was little socio-demographic variation in our study sample and no baseline characteristic predicted system use vs. non-use, but participants who reported greater perceived financial hardship used the system less often than those who felt more financially secure. This is despite the fact that those with greater perceived financial hardship actually rated the system as *more* flexible than subjects who felt more financially secure. This suggests that those who experience greater financial stress and its attendant life disruptions may particularly appreciate the flexibility and accessibility of computerized health promotion systems such as ours. Even though their life circumstances may make it difficult for them to fully utilize these systems, it is conceivable that their engagement in more rigid, traditional, face-to-face programs would be even lower. Similarly, we found that participants with lower educational attainment rated the system more favorably than more educated users in the areas of information content, applicability, and in their overall experience. Efforts were made to develop an intervention that was acceptable to a wide range of literacy levels. The system's ability to identify and address the user's knowledge gaps and to allow users to repeat conversations within modules may have enhanced its appeal among participants with less formal education. Further qualitative work is needed to identify barriers to engagement in automated behavior change programs.

Several additional issues may have reduced observed intervention effects. First, baseline values for several outcomes were unexpectedly high, which made it more difficult to achieve substantial improvements in the intervention group on these outcomes. Despite selection criteria designed to create a sample with uncontrolled BP and low adherence to self-care recommendations, mean baseline BP was only 131/81; 36% and 37% met recommendations for fruit and vegetable intake, respectively; and 32% met recommended guidelines for physical activity. In fact, 47% more participants in the intervention group met physical activity guidelines than in the control group (38.5% vs. 26.3%). This chance occurrence meant that a greater proportion of intervention group participants were advised to maintain rather than increase their levels of physical activity, and that downward regression to the mean for physical activity was more likely in this group. Additionally, since participants with low baseline physical activity were twice as likely to drop out of the intervention group than were control participants, the potential impact of upward regression to the mean was decreased in intent to treat analysis, also reducing physical-activity-related intervention effects.

The limited impact of the intervention on our three behavioral outcomes made it especially difficult to see an intervention effect on BP—a secondary outcome of this study—especially in the context of limited power. In our sample, there was less than 80% power to detect a difference in SBP of 7 mmHg between the experimental groups. The relationship between behavior change and resulting physiologic change is complex and depends on both the magnitude and duration of the observed behavioral changes. Further, BP reductions achieved by multi-behavior interventions are typically smaller than the sum of the reductions observed for their components when deployed alone. For example, the Trials of Hypertension Prevention Phase II study [27] showed an intervention combining weight loss and sodium restriction was associated with a smaller reduction in diastolic BP vs. controls at 6 months (−2.8 mmHg, $p<0.001$) than the sum of the two component interventions when deployed singly (−2.7 mmHg for weight loss alone, $p<0.001$, −1.6 for sodium alone, $p<0.001$). In the PREMIER trial [4], adding the original DASH diet to an "established" regimen consisting of weight loss, limited sodium and alcohol intake, and increased physical activity did not significantly reduce systolic BP further (−0.6 mmHg vs. "established", $p=0.43$).

Conclusion

This study is the first to evaluate a culturally adapted, computer-based telephone counseling system for African–Americans that targets three health-related behaviors for the self-management of hypertension. It was associated with modest improvements in several study outcomes, but

effects were hampered by high baseline attainment of goal levels for several principal study outcomes and by low system use. Nevertheless, computerized technologies like this one provide an opportunity to deliver behavior change programs on a public health scale at relatively low cost, and to tailor health messages to both ethnic/racial group and individual-level characteristics. Although our automated system was developed using input from African–Americans in Boston, it was designed to address obstacles that are likely to pose barriers to a broader segment of urban African–American of low socioeconomic status as well. As such, our results can be generalized to other urban settings in the USA. In addition, the content of the intervention can be further culturally adapted to reflect the cultural themes of specific African–American communities. More research is needed to better understand what combination of system features and user characteristics may lead to improved hypertension self-management among urban African–Americans of low SES.

Acknowledgments This study was funded by a grant from the National Heart, Lung, and Blood Institute (R01 HL69395). We would like to thank Kenneth Resnicow, PhD for his assistance in conceptualizing and designing this study; Lisa Marks, Adena Cohen-Barak, and Julia Gefter for their dedication to seeing this project through; and Amy Bachand, PhD for her assistance with the accelerometer data analysis.

Conflict of interest Dr. Friedman has stock ownership and a consulting agreement with InfoMedics, the company that owns commercial rights to the TLC technology used in the computerized intervention. He is also a member of its Board of Directors. None of the other authors has any potential conflicts of interest to disclose.

References

1. Ostchega Y, Yoon SS, Hughes J, Louis T. Hypertension awareness, treatment, and control—Continued disparities in adults: United States, 2005–2006. NCHS Data Brief No. 3. Hyattsville, MD: National Center for Health Statistics; 2008.
2. Appel L, Moore T, Obarzanek E, Vollmer W, Svetkey L, Sacks F, et al. A clinical trial of the effects of dietary patterns on blood pressure. New England Journal of Medicine 1997;336:1117–24.
3. Anonymous. Effects of weight loss and sodium reduction intervention on blood pressure and hypertension incidence in overweight people with high-normal blood pressure. The Trials of Hypertension Prevention, phase II. The Trials of Hypertension Prevention Collaborative Research Group. Archives of Internal Medicine. 1997;157:657–67.
4. Svetkey L, Harsha D, Vollmer W, Stevens V, Obarzanek E, Elmer P *et al.* Premier: A clinical trial of comprehensive lifestyle, Ann Epidemiol 2003;13:462–71.
5. Levkoff S, Sanchez H. Lessons learned about minority recruitment and retention from the Centers on Minority Aging and Health Promotion. Gerontologist 2003;43:18–26.
6. French S, Jeffery R, Story M, Neumark-Sztainer D. Perceived barriers to and incentives for participation in a weight-loss program among low-income women in WIC. J Am Diet Assoc. Jan 1998;98:79–81.
7. Noar S, Benac C, Harris M. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol. Bull. 2007;133:673–93.
8. Friedman R, Kazis L, Jette A, Smith M, Stollerman J, Torgerson J, Carey K. A telecommunications system for monitoring and counseling patients with hypertension. Impact on medication adherence and blood pressure control. Am J Hypertens. 1996;9:285–92.
9. Pinto B, Friedman R, Marcus B, Kelley H, Tennstedt S, Gillman M. Effects of a computer-based, telephone-counseling system on physical activity. American Journal of Preventive Medicine 2002;23:113–20.
10. King A, Friedman R, Marcus B, Castro, C, Napolitano M, Ahn D, Baker L. Ongoing physical activity advice by humans versus computers: The community health advice by telephone (CHAT) trial. Health Psychology. 2007; 26:718–727.
11. Delichatsios H, Friedman R, Glanz K, Tennstedt S, Smigelski C, Pinto B, et al. Randomized trial of a "talking computer" to improve adults' eating habits. Am J Health Promot 2001;15:215–24.
12. Pate R, Pratt M, Blair S, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995;273:402–7.
13. Morisky D, Alfonso A, Krousel-Wood M, Ward H. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens. 2008;10:348–54.
14. Bandura A. Social Learning Theory. Englewood Cliffs, NJ: Prentice Hall, 1977
15. Prochaska J, DiClemente C. Stage and processes of self-change in smoking: Towards an integrative model of change J Consult Clin Psych 1983;51:390–5.
16. Miller W, Rollnick S, editors. Motivational interviewing: Preparing people for change. New York: The Guilford Press; 2002. p. 284–98.
17. Resnicow K, Baranowski T, Ahluwalia J, Braithwaite R. Cultural sensitivity in public health: Defined and demystified. Ethnicity and Disease 1999;9:10–21.
18. Nobles W. African philosophy: Foundations for black psychology. In: Jones R, editor. Black Psychology, 2nd Ed. New York: Harper & Row; 1980. p. 23–36.
19. Blair SN, Haskell W, Ho P, Paffenbarger R, Jr., Vranizan K, Farquhar J, et al. Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. Am J Epidemiol 1985;122:794–804.
20. Sarkin J, Campbell L, Gross L, et al. Project GRAD seven-day physical activity recall interviewer's manual. Med Sci Sports Exerc 1997;29:S91–S102.
21. Yaroch A, Resnicow K, Davis M, Davis A, Smith M, Khan L. Development of a modified picture-sort food frequency questionnaire administered to low-income, overweight, African–American adolescent girls. J Am Diet Assoc 2000;100:1050–6.
22. Krauss R, Eckel R, Howard B, et. al. AHA Dietary Guidelines: Revision 2000: A statement for healthcare professionals from the Nutrition Committee of the American Heart Association. Circulation 2000;102;2284–99
23. The Trials of Hypertension Prevention Collaborative Research Group. The effects of nonpharmacologic interventions on blood pressure of persons with high normal levels. JAMA. 1992;267:1213–20.
24. Chobanian A, Bakris G, Black H, Cushman W, Green L, Izzo J, et al. The seventh report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure: The JNC 7 Report. JAMA 2003;289:2560–71.
25. Lachin JM. Statistical considerations in the intent-to-treat principle. *Controlled Clinical Trials* 2000; 21:167–189.

26. Stevens A., Migneault J., Fava J., Friedman R., Farzanfar R. Health technology questionnaire: Psychometric properties of a health behavior change measure. Annals of Behavioral Medicine. Annual meeting supplement, 2006; 31:S126.
27. The Trials of Hypertension Prevention Collaborative Research Group. Effects of weight loss and sodium reduction intervention on blood pressure and hypertension incidence in overweight people with high normal blood pressure: The trials of hypertension prevention, phase II. Arch Intern Med. 1997;157:657–67.
28. Krebs P, Prochaska J, Rossi J. A meta-analysis of computer-tailored interventions for health behavior change. Preventive Medicine 51;2010:214–21
29. Hyman D, Pavlik V, Taylor W, Goodrick G, Moye L. Simultaneous vs. sequential counseling for multiple behavior change. Arch Intern Med. 2007;167:1152–8.
30. Spring B, Pagoto S, Pingitore R, Doran N, Schneider K, Hedeker D. Randomized controlled trial for behavioral smoking and weight control treatment: Effect of concurrent versus sequential intervention. J Consult Clin Psychol. 2004;72:785–96.
31. Prochaska J. Multiple health behavior research represents the future of preventive medicine. Preventive Medicine. 2008;46:281–5
32. Artinian N, Fletcher G, Mozaffarian D, et. al. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: A scientific statement from the American Heart Association. Circulation. 2010;122:406–41
33. Sangha O, Stucki G, Liang M, Fossel A, Katz J. The self-administered co-morbidity questionnaire: A new method to assess co-morbidity for clinical and health services research. Arthritis Rheum. 2003;49:156–63.
34. Davis T, Long S, Jackson R, et al. Rapid estimate of adult literacy in medicine: A shortened screening instrument. Fam Med 1993;25:391–5.

Journal of Health Communication, 17:325–338, 2012

ISSN: 1081-0730 print/1087-0415 online
DOI: 10.1080/10810730.2012.715233

Routledge
Taylor & Francis Group

Health Literacy and 30-Day Postdischarge Hospital Utilization

SUZANNE E. MITCHELL, EKATERINA SADIKOVA, AND BRIAN W. JACK

Department of Family Medicine, Boston University School of Medicine/Boston Medical Center, Boston, Massachusetts, USA

MICHAEL K. PAASCHE-ORLOW

Section of General Internal Medicine, Boston University School of Medicine/Boston Medical Center, Boston, Massachusetts, USA

Low health literacy is associated with higher mortality, higher rates of hospitalization, and poor self-management skills for chronic disease. Early, unplanned hospital reutilization after discharge is a common and costly occurrence in U.S. hospitals. Still, few studies have examined the relation between health literacy and 30-day hospital reutilization rates. The authors examined the association between health literacy and 30-day reutilization of hospital services (readmission or return to the emergency department) in an urban safety net hospital, and conducted a secondary analysis of data from the control arm subjects of the Project RED and the RED-LIT trials. Health literacy was measured using the REALM tool. The primary outcome was rate of 30-day reutilization. The authors used multivariate Poisson regression analysis to control for potential confounding. Of the 703 subjects, 20% had low health literacy, 29% had marginal health literacy, and 51% had adequate health literacy. Sixty-two percent of subjects had a 12th-grade education or less. Subjects with low health literacy were more likely to be insured by Medicaid ($p < .001$); Black non-Hispanic ($p < .001$); unemployed, disabled, or retired ($p < .001$); low income ($p < .001$); and less educated (high school education or less, $p < .001$). The fully adjusted incidence rate ratio for low health literacy compared with adequate health literacy was 1.46 (CI [1.04, 2.05]). Low health literacy is a significant, independent, and modifiable risk factor for 30-day hospital reutilization after discharge. Interventions designed to reduce early, unplanned, hospital utilization after discharge should include activities to mitigate the effect of patients' low health literacy.

Nearly 20% of Medicare patients are readmitted to the hospital within 30 days of discharge (Jenks, Williams, & Coleman, 2009). Known predictors of early readmission include the following: lower socioeconomic status (Weissman, Stern, & Epstein, 1994), history of prior hospitalization (Van Walraven, Mamdani, Fang, & Austin, 2004) and advanced age (Cho, Lee, Arozullah, & Crittenden, 2008; Marcantonio et al., 1999),

This work was supported by the Agency for Health Research and Policy 1K08HS019771-01 (Dr. Mitchell). The authors acknowledge Ms. Jessica Martin for her support in data preparation.

Address correspondence to Suzanne E. Mitchell, Department of Family Medicine, Dowling 5, Boston Medical Center, 1 BMC Place, Boston, MA 02118, USA. E-mail: suzanne.mitchell@bmc.org

length of stay greater than 7 days (Krumholz et al., 1997), a high burden of comorbid illnesses (based on Charlson score; Librero, Peiro, & Ordinana, 1999), and specific diagnoses (e.g., depression, congestive heart failure, chronic obstructive pulmonary disease, myocardial infarction; Krumholz et al., 1997; Mitchell et al., 2010; Parashar et al., 2006; Wong et al., 2008). Meanwhile, a streamlined discharge process designed to enhance patient safety during care transitions can reduce 30-day postdischarge hospital utilization as much as 30% (Jack et al., 2009). The 30-day rehospitalization rate has emerged as an important indicator of hospital quality, particularly because the Centers for Medicare and Medicaid Services introduced a series of new reimbursement policies that include decreasing payments to hospitals with high rates of rehospitalization and bundled payment schemes in which accountable care organizations will be paid less for subsequent hospital utilization within 30 days. Thus, there is great interest in identifying modifiable risk factors for rehospitalization that could be used to refine intervention models and lead to improvements in quality of care, patient outcomes, and cost savings.

Low health literacy has been linked to poor health outcomes, particularly for patients with chronic conditions such as diabetes, asthma, cancer, depression (Lincoln et al., 2006), HIV/AIDS (Kalichman & Rompa, 2000), and heart failure (Peterson et al., 2011). *Health literacy*, defined as the degree to which individuals have the capacity to obtain, process, and understand health information, skills, and services needed to make informed health decisions and take informed actions, affects many Americans (Paasche-Orlow, 2011; U.S. Department of Health and Human Services, 2000). An estimated 26% of the U.S. population has low health literacy, and an additional 20% has marginal health literacy (Paasche-Orlow et al., 2005). Health literacy barriers are often associated with greater risk of hospitalization (Baker, Parker, Williams, & Clark, 1998), higher rates of self-reported poor health status (Baker, Parker, Williams, Clark, & Nurss, 1997) decreased knowledge of one's medical condition, poor medication recall, nonadherence to treatment plans, poor self-care behaviors (Evangelista et al., 2010), and increased all-cause mortality (Wolf, Feinglass, Thompson, & Baker, 2010). Patients with low health literacy also are more likely to report unsatisfactory patient–doctor communication at the time of discharge, suggesting that some physicians may be insensitive or unaware when their patients are having difficulties comprehending discharge instructions (Kripilani et al., 2010). Although several of these factors have been linked to an increased risk for 30-day rehospitalization, very limited evidence has been presented to demonstrate the independent association between low health literacy and hospital reutilization within 30 days of discharge.

This study examined the relation between health literacy and hospital reutilization within 30 days of discharge at the Boston Medical Center. The Boston Medical Center is the largest safety net hospital in New England, providing a spectrum of medical services to an urban, socially and economically diverse population. We hypothesized that low health literacy would be an independent risk factor for early unplanned hospital reutilization after discharge for general medical patients (i.e., adult patients admitted for acute general medical conditions such as pneumonia, unstable angina, pancreatitis, acute renal failure).

Method

We conducted a secondary analysis of the Project RED (Re-Engineered Discharge) and RED-Lit clinical trial data sets (clinicaltrials.gov identifier: NCT00252057) to assess the association between health literacy and the rate of subsequent 30-day hospital reutilization. The original Project RED included 738 participants and

RED-Lit included 802 participants. All Project RED studies enrolled English-speaking patients 18 years or older who were admitted to a general medical unit at the Boston Medical Center. Study subjects were required to have telephone access and be able to convey an understanding of study procedures and other consent elements in English. Participants were excluded if their admission was planned, they were on suicide watch, transferred from another health facility or were deaf or blind. Outcome data was not used in the secondary analysis if participants withdrew consent, died during the index admission, or were not discharged to the community. The combined sample included 1,540 patients from the control and intervention arms with complete information for the primary independent variable of interest—health literacy—and the outcome variable—30-day hospital utilization. We used a final sample of 703 patients identified from the control arms of each of these trials. Subjects from the intervention groups were excluded to eliminate effect modification introduced by exposure to the intervention.

Key Outcome Variables

The primary outcome variable for this analysis was a combined count measure of emergency department and hospital utilization events by a patient within a 30-day period after the index discharge. We also examined the emergency department revisit and hospital readmission count outcomes separately for the purposes of identifying factors associated with these distinct events. The number of utilizations ranged from 0 to 15; however, we top-coded to a count of 8 to avoid undue influence of outliers. We collected outcome data using the Boston Medical Center electronic medical record or participant self-report obtained by phone interview after 30 days. Both reutilizations of the Boston Medical Center and other hospitals and emergency departments were included.

Primary Independent Variable

We measured health Literacy using the Rapid Estimate of Adult Literacy in Medicine (REALM), which is a 66-item validated word recognition test (Davis et al., 1993). The REALM assigns a grade level of literacy, with scores of 0–18 corresponding to literacy of third grade or below, 19–44 to 4th–6th grade, 45–60 to 7th–8th grade, and 61–66 to 9th grade or above. For the purposes of our analysis, the two categories of lowest literacy were combined because of the distribution of scores. We administered the REALM in person to study participants.

Statistical Analysis

Descriptive Statistics

We performed bivariate analyses to assess the unadjusted relation between demographic and clinical characteristics and the three REALM categories of health literacy. The crosstabs reflect the results of chi-square tests. We then conducted Poisson regressions for all three outcome variables (combined reutilization, rehospitalization and return to the emergency department) to control for potential confounding. Predictors were included into the multivariable Poisson regressions on the basis of their effect on the association between health literacy and the outcomes, as well as the significance of their independent associations with the outcomes. Predictors were chosen from: age, gender,

marital status, income, insurance, employment, education, race, having a primary care physician, being homeless in the 6 months before the index admission, evidence of depressive symptoms, frequent utilizer status, length of stay, and medical comorbidity.

Adjusted Charlson Score for Comorbidity

Because of differences in the methods used to calculate Charlson scores between the RED and RED lit trials, we calculated a correction factor using a linear transformation procedure to adjust for the inconsistency in the distributions of Charlson scores for the combined analytic dataset. Charlson scores for the RED Lit II sample were calculated in a manner that accounted for comorbid conditions using all available records. For Project RED and RED Lit I, this process did not include outpatient record review. To derive a comparable Charlson score for the complete dataset, we calculated a z-score for each RED-lit II participant's observed score using the mean and standard deviation for the RED-lit II sample. Adjusted Charlson scores were then calculated for RED and RED-lit I (using the mean and standard deviation of this sample) to correspond to the *z* scores from the more comprehensively calculated sample.

Adjustment for Confounding

We constructed multivariable Poisson regression models using thorough backwards selection processes for each of the three outcome variables. We included certain variables (e.g., race, education, insurance) into the models because of their established relations with the variables of interest. We used SAS 9.1 to conduct the analysis with two-sided tests with $p < .05$ to judge significance.

Results

Of the 703 subjects, 138 (20%) patients had low health literacy (≤6th grade or REALM score of 0–18), 207 (29%) had marginal health literacy (7th to 8th grade or REALM score of 19–44) and 358 (51%) had adequate health literacy (≥9th grade or REALM score of 45–60). Study participants' mean age was 49.2 years of age, which did not differ by REALM score (see Table 1). Patients with low health literacy were more likely to be insured by Medicaid ($p < .001$); Black ($p < .01$), unemployed, disabled, or retired ($p < .001$); low income (<$40,000/year, $p < .001$); and less educated (high school education or less, $p < .001$). Of subjects, 29% ($n = 206$) were frequent utilizers, defined as two or more admissions in 6 months before index admission. The mean length of stay for the index admission was 2.8 days and did not differ significantly by REALM score. The mean Charlson Comorbidity Index score was 0.6, with a majority of the sample reporting a score of 0. There was no relation between comorbidity and health literacy.

The unadjusted 30-day postdischarge hospital reutilization incidence rate ratio for subjects with low health literacy compared to subjects with adequate health literacy was 1.76 (95% CI [1.21, 2.55]). After adjusting for potential confounding using a multivariate Poisson regression analysis—which included education, gender, marital status, income, race, affiliation with primary care provider, homelessness, depression, frequent utilizer status, age, length of stay, and the Charlson Comorbidity Index—we found that patients with low health literacy are 1.46 times (95% CI [1.04, 2.05]) more likely than patients with adequate health literacy to return to the hospital or emergency department within 30 days. Frequent utilizers were 2.04 times more likely

Table 1. Sociodemographic characteristics by health literacy status

	Health literacy status, by REALM* score				
	1 ≤6th-grade level ($n = 138$)	2 7th–8th-grade level ($n = 207$)	3 ≥9th-grade level ($n = 358$)	Total	p
Gender					
Female	63	113	196	372	.15
	45.7	54.6	54.9		
Male	75	94	161	330	
	54.4	45.4	45.1		
Total	138	207	357	702	
Marital status					
Single, never been married	63	101	175	339	.07
	45.7	48.8	48.9		
Divorced, separated, or widowed	36	56	104	196	
	26.1	27.1	29.1		
Unknown	7	1	5	13	
	5.1	0.5	1.4		
Married	32	49	74	155	
	23.2	23.7	20.7		
Total	138	207	358	703	
Income					
No income–$19,999	35	71	72	178	<.001
	28.0	35.9	20.6		
$20,000–$39,999	74	80	140	294	
	59.2	40.4	40.0		
$40,000–$74,999	12	30	70	112	
	9.6	15.2	20.0		

(*Continued*)

Table 1. Continued

	Health literacy status, by REALM* score				
	1 ≤6th-grade level ($n = 138$)	2 7th–8th-grade level ($n = 207$)	3 ≥9th-grade level ($n = 358$)	Total	p
Unknown or refused	0	7	22	29	
	0.0	3.5	6.3		
Total	125	198	350	673	
Insurance					
Free Care**	20	29	47	96	.02
	14.5	14.0	13.1		
Medicaid	70	97	133	300	
	50.7	46.9	37.2		
Medicare	24	21	35	80	
	17.4	10.1	9.8		
Other or unknown	8	10	30	48	
	5.8	4.8	8.4		
Private	16	50	113	179	
	11.6	24.2	31.6		
Total	138	207	358	703	
Employment					
Disabled or injured	46	53	74	173	.04
	35.1	25.7	21.0		
Part time	12	23	38	73	
	9.2	11.2	10.8		
Retired	24	23	54	101	
	18.3	11.8	15.3		
Student	2	4	9	15	
	1.5	1.9	2.6		

Unemployed	29	56	61	146	
	22.1	27.2	17.3		
Full time	18	47	116	181	
	13.7	22.8	33.0		
Total	131	206	352	689	
Education					
Incomplete high school	60	62	44	166	<.01
	43.5	30.0	12.3		
Complete high school	55	91	120	266	
	39.9	44.0	33.5		
Some college	9	33	106	148	
	6.5	15.9	29.6		
Unknown	6	2	3	11	
	4.4	1.0	0.8		
Complete college	8	19	85	112	
	5.8	9.2	23.7		
Total	138	207	358	703	
Race					
Black	83	134	152	369	<.01
	60.1	64.7	42.5		
Hispanic	21	23	34	78	
	15.2	11.1	9.5		
Other	7	9	16	32	
	5.1	4.4	4.5		
Unknown	9	10	9	28	
	6.5	4.8	2.5		
White	18	31	147	196	
	13.0	15.0	41.1		
Total	138	207	358	703	

(*Continued*)

Table 1. Continued

	Health literacy status, by REALM* score				
	1 ≤6th-grade level ($n = 138$)	2 7th–8th-grade level ($n = 207$)	3 ≥9th-grade level ($n = 358$)	Total	p
Has primary care physician					
No	28	38	57	123	.48
	20.3	18.4	15.9		
Yes	110	169	301	580	
	79.7	81.6	84.1		
Total	138	207	358	703	
Been homeless in the past 6 months					
Yes	22	24	40	86	.33
	15.9	11.7	11.2		
No	116	182	318	616	
	84.1	88.4	88.8		
Total	138	206	358	702	
Depressive symptoms***					
Depressive	40	45	78	163	.19
	29.0	21.7	21.8		
Not depressive	98	162	280	540	
	71.0	78.3	78.2		
Total	138	207	358	703	

Frequent utilizer status					
Two or more admissions in the 6 months before index admission	41 29.7	54 26.1	111 31.0	206	.46
Fewer than two admissions in the 6 months before the index admission	97 70.3	153 73.9	247 69.0	497	
Total	138	207	358	703	
Age, *M* (*SD*)	51.23 (13.91)	49.42 (13.51)	48.14 (14.77)	702	
Length of stay, *M* (*SD*)	2.69 (2.57)	2.86 (2.52)	2.72 (2.38)	702	
Charlson Comorbidity Index, *M* (*SD*)	0.719 (1.592)	0.595 (1.382)	0.499 (1.341)	703	.03

Note. Some columns may not add up to 100% because of omission of the "Other" categories. REALM: Rapid Estimate of Adult Literacy in Medicine.

*REALM scores are typically categorized into four categories; however, because of the distribution of the scores, the lowest two categories were combined into one category, resulting in three REALM categories.

**Free Care was a program in Massachusetts that aimed to fund medical care for uninsured individuals.

***Positive depression symptom screen determined by the Patient Health Questionnaire-9 screen tool, a nine-item 4-point Likert scale, standard scoring algorithm to screen for major and minor depression. A score of 5 or higher indicates a positive depression symptom screen.

than nonutilizers (incidence rate ratio = 2.04 [1.58, 2.62] $p < .0001$) and those who were homeless in the 6 months before their index admission were 1.54 times more likely than those who were not homeless (incidence rate ratio = 1.54 [1.14,2.08] $p = .0053$; see Table 2) to return to the hospital or emergency department. Female gender (incidence rate ratio = 0.75 [0.58, 0.96] $p = .02$) and Black race (incidence rate ratio = 0.73 [0.55, 0.98] $p = .04$) were both associated with lower risk for early hospital reutilization. In subanalyses of the outcome variable, we used multivariate Poisson regression to examine incidence rate ratios for emergency department visits and hospital readmissions separately. We found that, compared with patients with adequate health literacy, low health literacy patients were 1.71 times more likely to be readmitted to the emergency department ($p < .05$) and 1.67 (95% CI [0.98, 2.83], $p < .06$, times more likely to be readmitted into the hospital within 30 days of index admission.

Discussion

Concern about 30-day unplanned hospital reutilization is warranted as these events correlate with higher morbidity, mortality, and costs and is being used as a marker of the quality of hospital care (Ashton, Del Junco, Souchek, Wray, & Mansyur, 1997). Our study suggests that patients with low health literacy are more likely than patients with adequate health literacy to return to the hospital or emergency department within 30 days of discharge. These results were robust and health literacy remained an independent predictor of hospital utilization within 30 days of discharge after adjusting for a range of potential confounding phenomena, including education level. This study adds to growing evidence demonstrating the negative consequences of limited health literacy on patients' health and well-being (Davis & Wolf, 2004). It also establishes a link between health literacy and the process of care transitions, which has emerged as a centerpiece of efforts to improve quality and decrease cost.

There are several potential reasons as to why patients with low health literacy may return to the hospital soon after their discharge relative to those with higher health literacy levels. Health literacy may prevent patients from understanding their discharge instructions—including proper comprehension of their diagnosis and the treatment regimen (Anonymous, 2009; Williams et al., 1995). Williams and colleagues who showed that patients with low health literacy misread medication dosing (23.6% incorrect) and appointment slips (39.6% incorrect) in a survey of hospitalized patients while those with adequate literacy did well on these tasks. Patients may also have difficulty managing self-care instructions after hospital discharge, such as understanding symptoms of medication side effects and how to mitigate such phenomena, appreciating early signs of disease exacerbation and responding appropriately, and accessing and utilizing routine and urgent outpatient services. These issues may all contribute to how health literacy is related to returning to the hospital.

Our results bring attention to the need for interventions designed to decrease the rate of unplanned hospital reutilization to focus on health literacy. Such interventions can augment the level of patient and family education and guidance that is provided. Agency for Healthcare Research and Quality offers an evidence-based toolkit to address health literacy titled, "Health Literacy Universal Precautions Toolkit." The toolkit notably addresses four strategies for overcoming health literacy barriers. These include the following: (a) using easily understood spoken communication, (b) modifying written communication, (c) teaching self-management and empowerment, and (d) bolstering patient's support systems. Agency for Healthcare Research and

Table 2. Adjusted incident rate ratio of hospital utilization within 30 days of discharge

	Multivariate Poisson regression with reutilization			
	Estimate	Wald 95% confidence limits		*p*
Intercept	0.41	0.20	0.82	.01
REALM				
≤6th grade	1.46	1.04	2.05	.03
7th and 8th grade	1.36	1.00	1.85	.05
≥9th grade	1.00	1.00	1.00	
Education				
Incomplete high school	0.96	0.62	1.48	.85
High school or equivalent	1.18	0.80	1.72	.40
Incomplete college	0.87	0.55	1.36	.54
Unknown	0.65	0.14	3.13	.59
Complete college	1.00	1.00	1.00	
Gender				
Female	0.75	0.58	0.96	.02
Male	1.00	1.00	1.00	
Marital status				
Single, never married	0.86	0.63	1.19	.37
Divorced, separated, or widowed	1.05	0.74	1.49	.78
Unknown	0.76	0.21	2.72	.68
Married	1.00	1.00	1.00	
Insurance				
Free Care*	1.15	0.73	1.80	.55
Medicaid	1.24	0.89	1.74	.21
Medicare	1.04	0.64	1.68	.88
Other or unknown	0.96	0.49	1.91	.92
Private	1.00	1.00	1.00	
Race				
Black	0.73	0.55	0.98	.04
Hispanic	0.73	0.47	1.12	.14
Other	0.76	0.46	1.25	.28
White	1.00	1.00	1.00	
Has primary care physician				
No	0.96	0.69	1.34	.82
Yes	1.00	1.00	1.00	
Been homeless				
Yes	1.54	1.14	2.08	.005
No	1.00	1.00	1.00	
Depressive symptoms**				
Depressive	1.24	0.96	1.60	.09
Not depressive	1.00	1.00	1.00	
Frequent user				
Two or more admissions	2.04	1.58	2.62	<.0001

(*Continued*)

Table 2. Continued

	Multivariate Poisson regression with reutilization			
	Estimate	Wald 95% confidence limits		*p*
Fewer than two admissions	1.00	1.00	1.00	
Age, continuous, increment of 1 year	0.99	0.98	1.00	.14
Length of stay, continuous, increment of 1 day	0.99	0.94	1.04	.61
Charlson Comorbidity Index, continuous, increment of 1 unit	1.00	0.91	1.09	.92

Note. Some columns may not add up to 100% because of omission of the "Other" categories. REALM: Rapid Estimate of Adult Literacy in Medicine.

**Free Care was a program in Massachusetts that aimed to fund medical care for uninsured individuals.

**Positive depressive symptom screen determined by the Patient Health Questionnaire-9 screen tool, a nine-item 4-point Likert scale, standard scoring algorithm to screen for major and minor depression. A score of 5 or higher indicates a positive depression symptom screen.

Quality's approach advocates for creating an environment in which patients of all health literacy levels can thrive. Some of these methods include the following: drawing pictures, using plain (nonmedical) language, and using the teachback approach (Brach et al., 2012; Koh et al., 2012; Weiss, 2003).

Limitations

Major strengths of the present study are the reasonably large dataset and the broad range of covariates available for analyses. However, several limitations should be noted. First, this study was conducted using data from clinical trials implemented at a single safety net hospital; therefore, results may not be generalizable to other patient populations. Further, reutilization events outside of the Boston Medical Center were collected by subject self-report but were not independently confirmed. However, we were able to confirm 91% of all events by medical record review. Last, although we attempted to account for known confounders, other factors may also exist and could remain unaddressed.

Conclusion

Our study suggests that low literacy is significantly associated with a higher rate of 30-day postdischarge hospital utilization. Patient health literacy plays an influential role in health outcomes and low health literacy can be a significant barrier to patients' safe transitions from hospital to home. Future directions include interventions to improve patient education for care transitions, reduce the health literacy burden of the discharge process, and remove unnecessary complexity from critical self-care tasks.

References

Anonymous. (2009). Study shows readmissions drop when patients understand discharge instructions. *Hospital Case Management*, *17*, 71–73.

Ashton, C. M., Del Junco, D. J., Souchek, J., Wray, N. P., & Mansyur, C. L. (1997). The association between the quality of inpatient care and early readmission: A meta-analysis of the evidence. *Medical Care*, *35*, 1044–1059.

Baker, D. W., Parker, R. M., Williams, M. V., & Clark, W. S. (1998). Health literacy and the risk of hospital admission. *Journal of General Internal Medicine*, *13*, 791–798.

Baker, D. W., Parker, R. M., Williams, M. V., Clark, W. S., & Nurss, J. (1997). The relationship of patient reading ability to self-reported health and use of health services. *American Journal of Public Health*, *87*, 1027–1030.

Brach, C., Dreyer, B., Schyve, P., Hernandez, L. M., Baur, C., Lemerise, A. J, & Parker, R. (2012). *Attributes of a health literate organization*, Washington, DC: Institute of Medicine.

Cho, Y. I., Lee, S. Y., Arozullah, A. M., & Crittenden, K. S. (2008). Effects of health literacy on health status and health service utilization amongst the elderly. *Social Science and Medicine*, *66*, 1809–1816.

Davis, T. C., Long, S. W., Jackson, R. H., Mayeaux, E. J., George, R. B., Murphy, P. W., & Crouch, M. A. (1993). Rapid Estimate of Adult Literacy in Medicine: A shortened screening instrument. *Family Medicine*, *25*, 391–395.

Davis, T. C., & Wolf, M. S. (2004). Health literacy: Implications for family medicine. *Family Medicine*, *36*, 595–598.

Evangelista, L. S., Rasmusson, K. D., Laramee, A. S., Barr, J., Ammon, S. E., & Dunbar, S. Yancy, C. W. (2010). Health literacy and the patient with heart failure—Implications for patient care and research: A consensus statement of the Heart Failure Society of America. *Journal of Cardiac Failure*, *16*, 9–16.

Jack, B. W., Chetty, V. K., Anthony, D., Greenwald, J. L., Burniske, G. M., Johnson, A. E., … Culpepper, L. (2009). The reengineered hospital discharge program to decrease rehospitalization. *Annals of Internal Medicine*, *150*, 178–187.

Jenks, S. F., Williams, M. V., & Coleman, E. A. (2009). Rehospitalizations among patients in the Medicare fee-for-service program. *New England Journal of Medicine*, *360*, 1457–1459.

Kalichman, S. C., & Rompa D. (2000). Functional health literacy is associated with health status and health-related knowledge in people living with HIV-AIDS. *Journal of Acquired Immunodeficiency Syndrome*, *25*, 337–344.

Koh, H. K., Berwick, D. M., Clancy, C. M., Baur, C., Brach, C., Harris, L. M., & Zerhusen, E. G. (2012). New federal policy initiatives to boost health literacy can help the nation move beyond the cycle of costly 'crisis care.' *Health Affairs*, *31*, 434–443.

Kripalani, S., Jacobson, T. A., Mugalla, I. C., Cawthon, C. R., Niesner, K. J., & Vaccarino, V. (2010). Health literacy and quality of physician communication during hospitalization. *Journal of Hospital Medicine*, *5*, 269–275.

Krumholz, H. M., Parent, E. M., Tu, N., Vaccarino, V., Wang, Y., Radford, M. J., & Hennen, J. (1997). Readmission after hospitalization for congestive heart failure among Medicare beneficiaries. *Archives of Internal Medicine*, *157*, 99–104.

Librero, J., Peiro, S., & Ordinana, R. (1999). Chronic comorbidity and outcomes of hospital care: Length of stay, mortality and readmission at 30 and 365 days. *Journal of Clinical Epidemiology*, *52*, 171–179.

Lincoln, A., Paasche-Orlow, M. K., Cheng, D. M., Lloyd-Travaglini, C., Caruso, C., Saitz, R., & Samet, J. H. (2006). Impact of health literacy on depressive symptoms and mental health-related: Quality of life among adults with addiction. *Journal of General Internal Medicine*, *21*, 818–822.

Marcantonio, E. R., McKean, S., Goldfinger, M., Kleenfield, S., Yurkofsky, M., & Brennan, T. A. (1999). Factors associated with unplanned hospital readmission among patients 65 years of age and older in a Medicare managed care plan. *American Journal of Medicine*, *107*, 13–17.

Mitchell, S. E., Paasche-Orlow, M., Chetty, V. K., Forsythe, S., O'Donnell, J., Greenwald, J., & Jack, B. W. (2010). Post-discharge hospital utilization and depression symptoms among inpatient general medical patients. *Journal of Hospital Medicine*, *5*, 378–384.

Paasche-Orlow, M. (2011). Caring for patients with limited health literacy: A 76-year-old man with multiple medical problems. *JAMA*, *306*, 1122–1129.

Paasche-Orlow, M., Parker, R. M., Gazmararian, J. A., Nielsen-Bohlman, L. T., & Rudd, R. R. (2005). The prevalence of limited health literacy. *Journal of General Internal Medicine*, *20*, 175–184.

Parashar, S., Rumsfeld, J. S., Spertus, J. A., Reid, K. J., Wenger, N. K., Krumholz, H. M., & Vaccarino, V. (2006). Time course of depression and outcome of myocardial infarction. *Archives of Internal Medicine*, *166*, 2035–2043.

Peterson, P. N., Shetterly, S. M., Clarke, C. L., Bekelman, D. B., Chan, P. S., Allen, L. A., & Masoudi, F. A. (2011). Health literacy and outcomes among patients with heart failure. *JAMA*, *305*, 1695–701.

U.S. Department of Health and Human Services. (2000). *Healthy People 2010: Understanding and Improving Health with Understanding and Improving Health and Objectives for Improving Health.* Washington, DC: U.S. Government Printing Office.

Van Walraven, C., Mamdani, M., Fang, J., & Austin, P. C. (2004). Continuity of care and patient outcomes after hospital discharge. *Journal of General Internal Medicine*, *19*, 624–631.

Weiss, B. D. (2003). *Health Literacy: A Manual for Clinicians.* Chicago, IL: American Medical Association Foundation.

Weissman, J. S., Stern, R. S., & Epstein, A. M. (1994). The impact of patient socioeconomic status and other social factors on readmission: A prospective study in four Massachusetts hospitals. *Inquiry*, *31*, 163–172.

Williams, M. V., Parker, R. M., Baker, D. W., Parikh, N. S., Pitkin, K., Coates, W. C., Nurss, J. R. (1995). Inadequate functional health literacy among patients at two public hospitals. *JAMA*, *274*, 1677–1682.

Wolf, M. S., Feinglass, J., Thompson, J., & Baker, D. W. (2010). In search of 'low health literacy': Threshold vs. gradient effect of literacy on health status and mortality. *Social Science and Medicine*, *70*, 1335–1441.

Wong, A. W., Gan, W. Q., Burns, J., Sin, D. D., & van Eeden, S. F. (2008). Acute exacerbation of chronic obstructive pulmonary disease: influence of social factors in determining length of stay and readmission rates. *Canadian Respiratory Journal*, *15*, 361–364.

OPEN ACCESS Freely available online

PLOS ONE

Unhealthy Alcohol Use, HIV Infection and Risk of Liver Fibrosis in Drug Users with Hepatitis C

Roberto Muga[1]*, Arantza Sanvisens[1], Daniel Fuster[1,2], Jordi Tor[1], Elisenda Martínez[1], Santiago Pérez-Hoyos[3], Alvaro Muñoz[4]

1 Department of Internal Medicine, Hospital Universitari Germans Trias i Pujol, Universitat Autònoma Barcelona, Barcelona, Spain, **2** Section of General Internal Medicine, Boston Medical Center, Boston University School of Medicine, Boston, Massachusetts, United States of America, **3** Department of Public Health, Institut de Recerca Hospital Vall d'Hebrón, Universitat Autònoma Barcelona, Barcelona, Spain, **4** Department of Epidemiology, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, Maryland, United States of America

Abstract

Aim: To analyze alcohol use, clinical data and laboratory parameters that may affect FIB-4, an index for measuring liver fibrosis, in HCV-monoinfected and HCV/HIV-coinfected drug users.

Patients and Methods: Patients admitted for substance abuse treatment between 1994 and 2006 were studied. Socio-demographic data, alcohol and drug use characteristics and clinical variables were obtained through hospital records. Blood samples for biochemistry, liver function tests, CD4 cell count, and serology of HIV and HCV infection were collected at admission. Multivariate linear regression was used to analyze the predictors of FIB-4 increase.

Results: A total of 472 (83% M, 17% F) patients were eligible. The median age at admission was 31 years (Interquartile range (IQR) 27–35 years), and the median duration of drug use was 10 years (IQR 5.5–15 years). Unhealthy drinking (>50 grams/day) was reported in 32% of the patients. The FIB-4 scores were significantly greater in the HCV/HIV-coinfected patients (1.14, IQR 0.76–1.87) than in the HCV-monoinfected patients (0.75, IQR 0.56–1.11) ($p<0.001$). In the multivariate analysis, unhealthy drinking ($p=0.034$), lower total cholesterol ($p=0.042$), serum albumin ($p<0.001$), higher GGT ($p<0.001$) and a longer duration of addiction ($p=0.005$) were independently associated with higher FIB-4 scores in the HCV-monoinfected drug users. The effect of unhealthy drinking on FIB-4 scores disappeared in the HCV/HIV-coinfected patients, whereas lower serum albumin ($p<0.001$), a lower CD4 cell count ($p=0.006$), higher total bilirubin ($p<0.001$) and a longer drug addiction duration ($p<0.001$) were significantly associated with higher FIB-4 values.

Conclusions: Unhealthy alcohol use in the HCV-monoinfected patients and HIV-related immunodeficiency in the HCV/HIV-coinfected patients are important risk factors associated with liver fibrosis in the respective populations

Citation: Muga R, Sanvisens A, Fuster D, Tor J, Martínez E, et al. (2012) Unhealthy Alcohol Use, HIV Infection and Risk of Liver Fibrosis in Drug Users with Hepatitis C. PLoS ONE 7(10): e46810. doi:10.1371/journal.pone.0046810

Editor: Michael Alan Polis, National Institute of Allergy and Infectious Diseases, United States of America

Received May 2, 2012; **Accepted** September 6, 2012; **Published** October 9, 2012

Funding: This work was funded by grants from Ministry of Science and Innovation, Spain (grants RD06/001, RD06/006 and PI05/1550), the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) (grant 2008 BE-2 00269), Ministry of Health (grant EC11-042) and Ministry of Education (grant 2010-0945). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rmuga.germanstrias@gencat.cat

Introduction

Liver fibrosis is the main predictor of whether chronic hepatitis C will progress to cirrhosis and end-stage liver disease [1]. Because the complications of liver disease mainly occur in patients with advanced-stage fibrosis, assessing chronic hepatitis C early is essential when evaluating at-risk patients [2]. In Western countries, more than 50% of new HCV infections are associated with drug abuse. However, this particular population also has lower rates of clinical assessment and chronic hepatitis C treatment. Given the likelihood of new and more effective treatments, drug abusers with chronic hepatitis C would benefit from simple, non-invasive measurements of liver fibrosis.

The cofactors associated with chronic hepatitis C progression differ among studies; alcohol abuse, male gender, age at infection, body mass index, and coinfection with human immunodeficiency virus infection (HIV) and Hepatitis B virus infection (HBV) have been related to more rapid disease progression [1–5]. In HCV/HIV-coinfected individuals, CD4 cell counts below 200 cells/μL have been associated with liver fibrosis progression [6]. In parallel, highly active antiretroviral therapy (HAART) has been shown to reduce liver-related deaths [7,8].

In HIV-negative patients, it is well established that alcohol abuse and HCV infection have a synergistic effect on liver fibrosis. However, there are conflicting results regarding the independent effect of alcohol on liver damage in HCV/HIV-coinfected patients [6,9,10].

Liver biopsy is the gold standard for assessing fibrosis [11]. However, assessing liver disease through an invasive procedure is unlikely in patients with substance abuse [12]. Furthermore, eligibility for chronic hepatitis C treatment in this population is low compared with eligibility in other populations [13,14]. To

a certain extent, the evolution of liver disease in drug abusers parallels the natural history of chronic hepatitis C.

Several non-invasive markers of liver fibrosis have been proposed as alternatives to liver biopsy. Some of these markers reflect the modified extracellular matrix turnover that occurs during fibrogenesis [15,16], whereas others reflect alterations in hepatic function [17,18]. FIB-4 was initially described in 2006 [18], and since then, it has been proposed as reliable marker of fibrosis in both HCV-monoinfected and HCV/HIV-coinfected individuals [18,19]. FIB-4 correlates well with liver biopsy in patients with and without advanced fibrosis [20,21]. Moreover, non-invasive markers of liver fibrosis have been proposed as predictors of all-cause and liver-related mortality [22,23].

Although abuse of alcohol and illegal drugs is frequent in patients with HIV infection and HCV infection, it is unclear how non-invasive liver fibrosis tests may reflect disease progression. In this study, we hypothesize that certain clinical and laboratory characteristics may influence a simple index of fibrosis and that the cofactors associated with elevated FIB-4 scores may differ between HCV-monoinfected patients and HCV/HIV-coinfected patients. Hence, the primary objective of the study was to characterize the putative differences in risk factors for elevated liver function biomarkers between HCV-monoinfected and HCV/HIV-coinfected patients.

Patients and Methods

Study Population

This was a cross-sectional study of patients admitted for substance abuse treatment between 1994 and 2006. The demographic and drug use characteristics were recorded through a structured questionnaire administered by a physician the day of admission. Questions related to drug and alcohol abuse included: (i) the main drug of abuse (type of drug, age at first use, duration of drug use and route of administration), (ii) poly-drug use (yes/no) (iii) alcohol consumption: do you regularly drink alcohol? (yes/no); if yes, do you drink 5 or more standard drinks per day?. A standard drink unit contains 12–14 grams of alcohol and unhealthy alcohol consumption was defined as a daily alcohol intake $\geq$50 grams (g) [24,25] in the 6-month period prior to admission. All participants gave written informed consent. The methods used in this study complied with the ethical standards for medical research and principles of good clinical practice defined by the World Medical Association's Declaration of Helsinki. The study was approved by the Ethics Committee of the Hospital Universitari Germans Trias i Pujol.

Routine laboratory parameters, including liver function tests and serology for HIV infection and HCV infection, were analyzed at admission. Other characteristics of admission for substance abuse treatment have been described elsewhere [26].

The liver function tests and biochemical parameters were assessed using an Olympus 5200 Multichannel chemistry analyzer. The procedure, which remained the same throughout the study, was based on the reference method recommended by the International Federation of Clinical Chemistry.

HIV infection was identified by an enzyme-linked immunosorbent assay. Repeatedly reactive samples were confirmed by the Western immunoblot technique.

HCV infection was assessed prior to or during admission by a second- or later-generation enzyme immunoassay (Ortho Diagnostics, Raritan, NJ). The positive samples were confirmed by either a recombinant immunoblot assay (RIBA HCV 2 SIA, Chiron Corporation, Emeryville, CA) or a qualitative/quantitative assay (COBAS AMPLICOR, Roche Diagnostic Systems, Branchburg, NJ).

Outcome

The primary outcome was the FIB-4 score, which was calculated as

$$FIB4 = \frac{Age[years] \times AST[U/L]}{Platelet[10^9/L] \times \sqrt{ALT[U/L]}}$$

FIB-4 scores lower than 1.45 indicate lack of liver fibrosis with a negative predictive value of 90% and a sensitivity of 70% [18]. FIB-4 scores greater than 3.25 indicate significant liver fibrosis with a positive predictive value of 65% and a specificity of 97% [18].

Statistical Analysis

All of the analyses were conducted separately for the HCV-monoinfected (N = 228) and the HCV/HIV-coinfected (N = 244) individuals. We used medians and interquartile ranges (IQRs) to describe the quantitative variables and absolute frequencies and percentages to describe the qualitative variables.

The distribution of FIB-4 score was strongly skewed to the right (i.e., there were several very high values); we therefore normalized it for analysis purposes using a logarithmic transformation.

We used multiple linear regression models to determine the FIB-4 predictive values of the variables. There were three types of predictors: (1) binary, which included sex, alcohol use, and HBsAg; (2) continuous on a natural (additive) scale, which included body mass index (BMI), CD4 cell count, total cholesterol, alkaline phosphatase, and duration of drug use; and (3) continuous on a logarithmic (multiplicative) scale, which included total bilirubin, serum albumin, and GGT. The decision to analyze a variable using a logarithmic scale was based on the need to reduce the undue influence of high values in predictors with strong right skewness.

The interpretation of the regression coefficients differed among the three types of predictors. Specifically, the regression coefficients of the binary variables represented the percentage FIB-4 difference between those with and without the condition; the regression coefficients of the additive continuous variables represented the percentage FIB-4 difference associated with an unitary increase or decrease in the variables, and the regression coefficients of the multiplicative continuous variables represented the percentage FIB-4 difference associated with an increment or decrement in the variables.

The intercept represented the expected FIB-4 score in an individual with zero values for all of the predictors.

The test results were considered to be statistically significant if the resulting P-*value* was $<.05$. The statistical analysis was performed using the SPSS software, version 15.0.1 (SPSS, Chicago, IL, USA).

Results

Patients were eligible for this study if they had chronic HCV infection (N = 544). Patients with aminotransferase levels 10 times greater than the upper limit of the normal range (N = 5,1.0%), patients who had received HCV antiviral therapy (N = 6,1.1%) and patients with antecedent of decompensated liver cirrhosis (N = 10,1.8%) were excluded. In addition, patients with an HCV-RNA level below the limit of detection (<50 IU/mL) were

excluded (N = 9, 1.7%). Finally, patients with outlier laboratory values and those with incomplete data for calculating FIB-4 score were also excluded (N = 42, 7.7%). After these exclusions, the study population consisted of 472 patients and 244 patients (52%) were coinfected with HIV. Table 1 shows the descriptive statistics at admission for the entire group and for the HCV-monoinfected (N = 228) and HCV/HIV-coinfected (N = 244) subgroups. Overall, 17% of the patients were women, the median age at admission was 31 years (IQR 27–35 years), the median BMI was 22 kg/m^2, the median duration of drug use was 10 years, and unhealthy drinking was reported in 32% of the patients. In addition to decreased CD4 cell counts, the HCV/HIV-coinfected patients had a longer median drug use duration, an increased frequency of unhealthy alcohol intake, lower levels of total cholesterol, higher levels of GGT and a higher prevalence of HBsAg than those infected with HCV only. The median AST and ALT levels and platelet counts were 35 U/L, 47 U/L, and 180 × 10^9/L, respectively.

Thirty-one percent of the HIV-positive patients were receiving antiretroviral therapy at admission, and 48% had never received antiretroviral therapy.

The median FIB-4 score at admission was 0.93 (IQR 0.65–1.46); it was significantly higher in the HCV/HIV-coinfected patients (1.14, IQR 0.76–1.87) than in the HCV-monoinfected patients (0.75, IQR 0.56–1.11). Figure 1 shows the distribution of the FIB-4 scores in the two groups on both natural and logarithmic scales. As can be seen in the bottom panels of Figure 1, the log-transformed FIB-4 scores approximately followed a normal distribution, rendering normally based methods appropriate for the analysis.

Regression Analysis of FIB-4 (Log Scale)

We conducted univariate regressions of the variables shown in Table 1 against the log-transformed FIB-4 scores; only the variables that define FIB-4 were not used in the univariate regressions. The two columns with univariate headings in Table 2 show the results of the univariate analyses separately for the monoinfected and coinfected patients.

In the univariate models for the HCV-monoinfected patients, unhealthy alcohol use, higher BMI, longer duration of drug use, lower cholesterol, higher bilirubin, lower albumin, and higher GGT were significantly associated ($p<0.05$) with higher FIB-4 scores. In the coinfected patients, higher FIB-4 scores were found to be significantly associated ($p<0.05$) with a longer drug use duration, lower cholesterol, higher alkaline phosphatase, lower CD4 cell count, higher bilirubin, lower albumin, and higher GGT.

For each group, the variables that showed significant relationships in the univariate analyses were used in a multivariate model. In the multivariate model for the HCV-monoinfected patients, unhealthy alcohol use ($p = 0.034$), longer drug use duration ($p = 0.005$), lower cholesterol ($p = 0.042$), lower albumin ($p<0.001$), and higher GGT ($p = 0.001$) continued to be significantly associated with higher FIB-4 scores.

In the coinfected patients, longer drug use duration ($p<0.001$), lower CD4 cell count ($p = 0.007$), higher bilirubin ($p<0.001$), and lower albumin ($p<0.001$) were significantly associated with higher FIB-4 scores.

Longer drug use duration and lower albumin levels were significantly associated with increased FIB-4 scores in both groups. By contrast, unhealthy alcohol use was strongly predictive of high FIB-4 scores only in the HCV-monoinfected group; similarly, high

Table 1. Descriptive statistics (median [IQR] or n (%)) of HCV-monoinfected and HCV/HIV-coinfected patients at admission to substance abuse treatment.

	HCV	HCV/HIV		Total
	N = 228	**N = 244**	**p_value***	**N = 472**
Socio-demographic and anthropometric				
Females	35 (15.4%)	45 (18.4%)	0.371	80 (16.9%)
Age, years	30 [26,34]	31.5 [28, 35]	0.008	31 [27, 35]
Body mass index, kg/m^2 (N = 421)	22.3 [20.7, 24.6]	21.6 [19.6, 23.6]	0.000	21.9 [20.2, 23.9]
Drug use				
Unhealthy alcohol use (N = 440)	62 (28.2%)	79 (35.9%)	0.082	141 (32.0%)
Duration of drug use, years (N = 463)	7.6 [3.5, 12.0]	12.0 [7.6, 16.0]	0.000	10.0 [5.5, 15.0]
Laboratory parameters				
Hepatitis B surface Antigen (N = 440)	9 (4.2%)	15 (6.7%)	0.243	24 (5.5%)
Total cholesterol, mg/dL	162 [143, 178]	149 [128, 170]	0.001	155 [135, 174]
Alkaline Phosphatase, U/L	70 [55, 82]	70 [59, 86]	2.664	70 [56, 84]
Total bilirubin, mg/dL	0.4 [0.3, 0.6]	0.4 [0.3, 0.6]	0.297	0.4 [0.3, 0.6]
Albumin, g/L (N = 438)	39 [36, 42]	38 [35, 41]	0.013	39 [36, 41]
GGT, U/L	33 [19, 62]	44.5 [24, 93]	0.052	38 [22, 76]
CD4 lymphocytes, cells/μL (N = 456)	1225 [933, 1428]	383 [204, 661]	0.000	742 [350, 1225]
Laboratory components of FIB-4				
Platelets, 10^9/L	197 [166, 243]	163 [127, 196]	0.000	180 [146, 224]
AST, U/L	33.5 [21.0, 61.0]	37.5 [24.2, 61.0]	0.907	35.0 [23.0, 61.0]
ALT, U/L	54.0 [23.0, 98.0]	43.0 [25.0, 71.0]	0.001	47.0 [24.0, 84.0]

*p value for the comparison between HCV-monoinfected and HCV/HIV-coinfected patients; p values correspond to χ square test in categorical variables and t test for differences of mean values in continuous variables.

doi:10.1371/journal.pone.0046810.t001

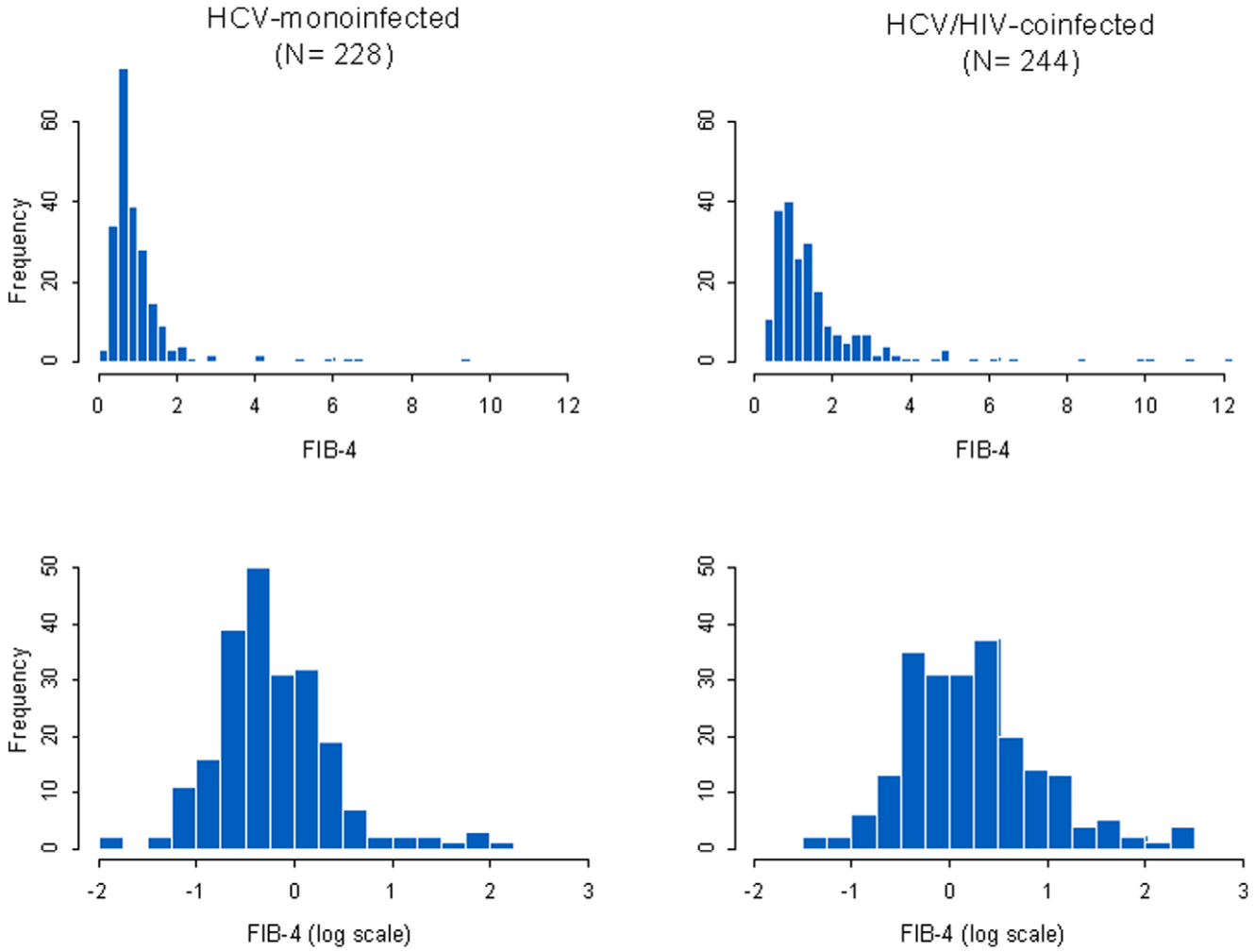


Figure 1. Distribution of FIB-4 score and log FIB-4 score according to HCV-monoinfection and HCV/HIV-coinfection.
doi:10.1371/journal.pone.0046810.g001

total bilirubin levels were associated with higher FIB-4 scores only in the coinfected patients.

Although the primary aim of the study was the characterization of the variables that were associated with FIB-4 increase in HCV-monoinfected and HCV/HIV-coinfected, the intercepts of the multivariate models provide a means of comparing a hypothetical HCV-monoinfected individual with a hypothetical HCV/HIV-coinfected individual, assuming that both have 900 CD4 cells/μL and that all of the other variables are equal. The slightly increased FIB-4 intercept value (0.778 in the monoinfected and 0.875 in the coinfected patients) was not statistically significant ($p = 0.218$). However, for each decline of 100 CD4 cells/μL among the coinfected patients, there was a significant FIB-4 increase of 3.6% ($p = 0.007$) (Table 2).

To further characterize the differential effect of unhealthy alcohol use on FIB-4 scores, Figure 2 shows the distributions of the FIB-4 scores in the four groups by alcohol consumption and HIV coinfection. Figure 2 shows box plots that are graphically enhanced to show the 2.5^{th}, 5^{th}, 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, 95^{th} and 97.5^{th} percentiles of the FIB-4 distribution. It is clear from Figure 2 that the primary difference was between the non-drinking HCV-monoinfected patients and the other three groups. In particular, the increased FIB-4 score due to unhealthy alcohol use among the HCV-monoinfected patients was similar to the effect of HIV-related immunodeficiency among the non-drinkers. The additional increase in FIB-4 among the coinfected individuals with unhealthy alcohol use was small and not significant ($p = 0.695$).

Discussion

Individuals with histories of drug use account for the majority of new hepatitis C infections in Western countries. This population is at risk for liver fibrosis, and a number of disease progression cofactors highlight the relevance of medical assessment. Evaluating liver fibrosis via non-invasive tests early in the course of drug addiction may increase the proportion of patients who are eligible for treatment. This study of young adults with chronic hepatitis C shows that the factors associated with higher FIB-4 scores clearly differed between the HCV-monoinfected and HCV/HIV-coinfected individuals.

The main contribution of the study is related to the fact that unhealthy alcohol use had a differential effect on FIB-4 values if patients have hepatitis C alone or HCV/HIV coinfection. Unhealthy alcohol drinking has been regarded as a major contributor to the progression of liver disease in the setting of chronic hepatitis C [27] and, a synergistic effect between HCV and alcohol has been proposed [28]. However, even though

Table 2. Percentage change in FIB-4 score associated with differences in predictors of higher FIB-4 scores.

	HCV-monoinfected		HCV/HIV-coinfected	
	N = 228		N = 244	
	Univariate	Multivariate	Univariate	Multivariate
Variable	**% (p_value)**	**% (p_value)**	**% (p_value)**	**% (p_value)**
Intercept	NA	0.778*	NA	0.875**
		(95% CI : 0.705, 0.861)		(95% CI: 0.762, 1.005)
Female to Male	−11.6% (0.266)		+11.5% (0.343)	
Unhealthy alcohol use	**+46.7% (<.001)**	**+20.6% (0.034)**	+3.9% (0.695)	
HBsAg positive	+25.1% (0.268)		+31.3% (0.155)	
BMI: increase of 1 kg/m^2	**+4.3% (0.001)**	+2.2% (0.069)	+0.8% (0.609)	
Duration of drug use: increase of 5 years	**+9.4% (0.007)**	**+10.0% (0.005)**	**+21.5% (<.001)**	**+13.9% (<.001)**
Cholesterol: decrease of 20 mg/100 mL	**+5.4% (0.018)**	**+4.5% (0.042)**	**+6.6% (0.006)**	+2.4% (0.318)
Alkaline Phosphatase: increase of 10 U/L	+2.7% (0.108)		**+4.9% (0.001)**	+0.9% (0.618)
CD4: decrease of 100 cells/μL	NA	NA	**+5.2% (<.001)**	**+3.6% (0.007)**
Bilirubin: 1.5-fold increment	**+12.7% (<.001)**	+5.7% (0.097)	**+22.8% (<.001)**	**+22.2% (<.001)**
Albumin: 1.1-fold decrement	**+19.8% (<.001)**	**+19.1% (<.001)**	**+15.4% (<.001)**	**+12.9% (<.001)**
GGT: 2-fold increment	**+20.3% (<.001)**	**+12.3% (0.001)**	**+16.4% (<.001)**	+7.3% (0.061)

*Expected value of FIB-4 for individuals with predictors at BMI = 22 kg/m^2, no alcohol consumption, total cholesterol = 155 mg/100 mL, Total bilirubin = 0.4 mg/dL, albumin = 39 g/L, GGT = 38 U/L, and duration of IDU = 10 years.
**Expected value of FIB-4 for individuals with predictors at CD4 = 900 cells/μL, total cholesterol = 155 mg/100 mL, total bilirubin = 0.4 mg/dL, albumin = 39 g/L, Alkaline Phosphatase = 70 U/L, GGT = 38 U/L, and duration of IDU = 10 years.
doi:10.1371/journal.pone.0046810.t002

alcohol use is among the cofactors related with liver fibrosis in coinfection in studies with liver biopsy [6,29], the present and other studies that have used non- invasive methods to estimate fibrosis [9,10], does not detect an additional effect of alcohol drinking on the FIB-4 of HCV/HIV-coinfected patients.

In coinfected patients with unhealthy alcohol consumption, the FIB-4 does not reflect the negative impact of alcohol intake on liver fibrosis. Therefore, clinicians may not be able to assess the impact of ethanol nor can advise the patient on the risk of disease progression. On the contrary, unhealthy alcohol use is reflected in

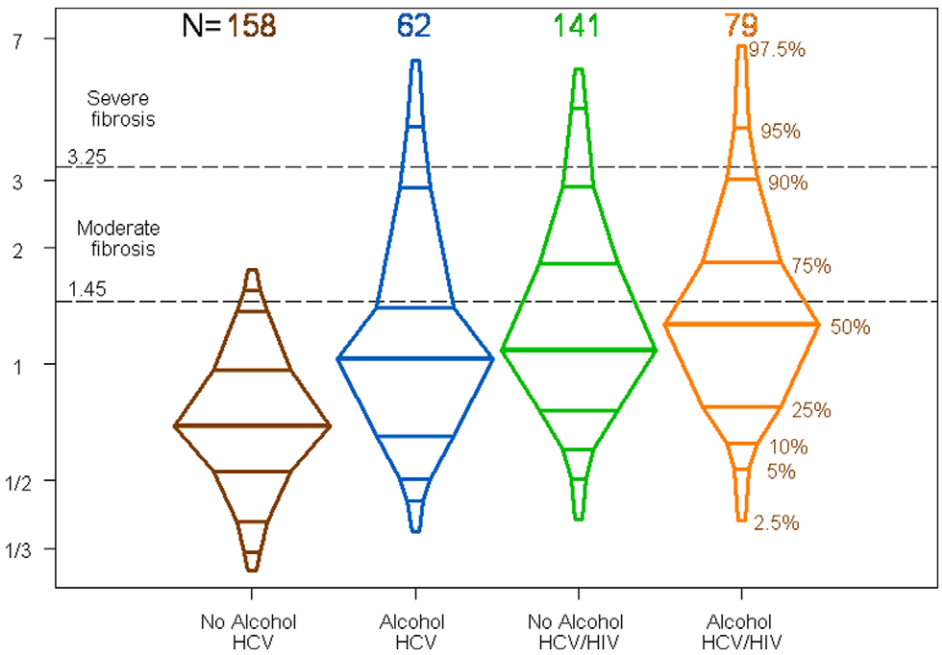


Figure 2. Distribution of FIB-4 scores according to unhealthy alcohol use in HCV-monoinfected and HCV/HIV-coinfected patients.
doi:10.1371/journal.pone.0046810.g002

the FIB-4 of the monoinfected patients thus making possible preventive interventions to reduce harm.

Unhealthy alcohol use in the HCV-monoinfected patients and HIV-related immunodeficiency in the HCV/HIV-coinfected patients are the most important cofactors associated with fibrosis progression in the respective populations. In addition, we found that drug use duration and serum albumin were correlated with the FIB-4 scores of both the monoinfected and coinfected patients, whereas unhealthy alcohol use, GGT and total cholesterol were associated with higher FIB-4 scores only in the monoinfected patients. The effect of HIV-related immunodeficiency in the coinfected patients was strong (an increase of 3.5% in the FIB-4 score for every 100 CD4 cells/μL decrease). Furthermore, we did not observe differing FIB-4 values between the HCV-monoinfected and coinfected individuals with CD4 cell counts above 900 cells/μL. This observation suggests that FIB-4 elevation is associated with immunoactivation and the resulting decrease of CD4 cell counts in HCV/HIV-coinfected drug users.

The relationship between HIV-related immunodeficiency and liver fibrosis progression has been described in coinfected patients [6,29,30]; in fact, treating HIV/AIDS with HAART has been shown to reverse the effect of HIV-related immunodeficiency in patients with chronic hepatitis C [7,26,29,31].

In this study, decreased serum albumin and increased total bilirubin were associated with elevated FIB-4 scores. This finding may facilitate identifying a subpopulation of patients at increased risk for cirrhosis. It is well known that albumin and bilirubin are key components of the Child-Turcotte-Pugh score that clinicians use to assess decompensated liver cirrhosis.

In individuals with history of injection drug use, the duration of injection use is a surrogate for the duration of HCV infection [32]. As expected, the duration of drug addiction in this study was related to increased FIB-4 scores.

It has been reported that HCV infection itself lowers both low-density lipoprotein (LDL) and total cholesterol and that patients treated for chronic hepatitis C had larger increases in LDL and total cholesterol from baseline [33]. Interestingly, the current study did not find a significant association between cholesterol levels and FIB-4 scores among the HCV/HIV-coinfected patients.

This study has a number of limitations that should be mentioned. First, the alcohol intake assessment was limited to one categorical variable (>50 g/day, ≤50 g/day in the 6-month period before admission), and there was no information on the history and complications of alcohol consumption. In previous studies, however, recent alcohol consumption has been treated as a dichotomous variable using a threshold of 40–50 grams of ethanol per day or using the definition of heavy alcohol intake provided by the US National Institute on Alcohol Abuse and Alcoholism (NIAAA) [6,9,10]. Second, we used a single measurement of laboratory parameters to calculate the FIB-4 scores which precluded examining the evolution of fibrosis over time. Furthermore, nearly half of the patients had normal aminotransferases values, as has been previously described for IDUs [34]; despite the lack of correlation between liver enzyme alterations and liver damage, it is possible that FIB-4 scores are affected by normal liver enzymes [20]. Third, the HAART status of the HIV-positive patients was represented by a qualitative covariate, which hindered an analysis of the effect of antiretroviral treatment on FIB-4.

In summary, this study shows that unhealthy alcohol use strongly influence FIB-4 in HCV- monoinfected patients, whereas in the context of HCV/HIV coinfection, HIV-related immune depression exerts a major negative role on FIB-4 results, with no significant worsening by alcohol intake.

Author Contributions

Conceived and designed the experiments: RM AM. Performed the experiments: RM JT DF. Analyzed the data: AM AS SPH EM. Contributed reagents/materials/analysis tools: AS. Wrote the paper: RM AM DF.

References

1. Thomas DL, Seeff LB (2005) Natural history of hepatitis C. Clin Liver Dis 9: 383,98, vi DOI:10.1016/j.cld.2005.05.003.
2. Massard J, Ratziu V, Thabut D, Moussalli J, Lebray P, et al. (2006) Natural history and predictors of disease severity in chronic hepatitis C. J Hepatol 44: S19–S24 DOI:10.1016/j.jhep.2005.11.009.
3. Davis GL, Lau JY (1997) Factors predictive of a beneficial response to therapy of hepatitis C. Hepatology 26: 122S–127S DOI:10.1002/hep.510260721.
4. Poynard T, Ratziu V, Charlotte F, Goodman Z, McHutchison J, et al. (2001) Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis c. J Hepatol 34: 730–739.
5. Sulkowski MS, Thomas DL (2003) Hepatitis C in the HIV-Infected Person. Ann Intern Med 138: 197–207.
6. Benhamou Y, Bochet M, Di Martino V, Charlotte F, Azria F, et al. (1999) Liver fibrosis progression in human immunodeficiency virus and hepatitis C virus coinfected patients. The Multivirc Group. Hepatology 30: 1054–1058 DOI:10.1002/hep.510300409.
7. Benhamou Y, Di Martino V, Bochet M, Colombet G, Thibault V, et al. (2001) Factors affecting liver fibrosis in human immunodeficiency virus-and hepatitis C virus-coinfected patients: impact of protease inhibitor therapy. Hepatology 34: 283–287 DOI:10.1053/jhep.2001.26517.
8. Qurishi N, Kreuzberg C, Luchters G, Effenberger W, Kupfer B, et al. (2003) Effect of antiretroviral therapy on liver-related mortality in patients with HIV and hepatitis C virus coinfection. Lancet 362: 1708–1713 DOI:10.1016/S0140-6736(03)14844-1.
9. Blackard JT, Welge JA, Taylor LE, Mayer KH, Klein RS, et al. (2011) HIV mono-infection is associated with FIB-4 - A noninvasive index of liver fibrosis - in women. Clin Infect Dis 52: 674–680 DOI:10.1093/cid/ciq199.
10. Chaudhry AA, Sulkowski MS, Chander G, Moore RD (2009) Hazardous drinking is associated with an elevated aspartate aminotransferase to platelet ratio index in an urban HIV-infected clinical cohort. HIV Med 10: 133–142 DOI:10.1111/j.1468-1293.2008.00662.x.
11. Gebo KA, Herlong HF, Torbenson MS, Jenckes MW, Chander G, et al. (2002) Role of liver biopsy in management of chronic hepatitis C: a systematic review. Hepatology 36: S161–S172 DOI:10.1053/jhep.2002.36989.
12. Sulkowski MS, Thomas DL (2005) Epidemiology and natural history of hepatitis C virus infection in injection drug users: implications for treatment. Clin Infect Dis 40 Suppl 5: S263–S269 DOI:10.1086/427440.
13. Fleming CA, Craven DE, Thornton D, Tumilty S, Nunes D (2003) Hepatitis C virus and human immunodeficiency virus coinfection in an urban population: low eligibility for interferon treatment. Clin Infect Dis 36: 97–100 DOI:10.1086/344907.
14. Zinkernagel AS, von Wyl V, Ledergerber B, Rickenbach M, Furrer H, et al. (2006) Eligibility for and outcome of hepatitis C treatment of HIV-coinfected individuals in clinical practice: the Swiss HIV cohort study. Antivir Ther 11: 131–142.
15. Leroy V, Monier F, Bottari S, Trocme C, Sturm N, et al. (2004) Circulating matrix metalloproteinases 1, 2, 9 and their inhibitors TIMP-1 and TIMP-2 as serum markers of liver fibrosis in patients with chronic hepatitis C: comparison with PIIINP and hyaluronic acid. Am J Gastroenterol 99: 271–279.
16. Sanvisens A, Serra I, Tural C, Tor J, Ojanguren I, et al. (2009) Hyaluronic acid, transforming growth factor-beta1 and hepatic fibrosis in patients with chronic hepatitis C virus and human immunodeficiency virus co-infection. J Viral Hepat 16: 513–518 DOI:10.1111/j.1365-2893.2009.01103.x.
17. Pohl A, Behling C, Oliver D, Kilani M, Monson P, et al. (2001) Serum aminotransferase levels and platelet counts as predictors of degree of fibrosis in chronic hepatitis C virus infection. Am J Gastroenterol 96: 3142–3146 DOI:10.1111/j.1572-0241.2001.05268.x.
18. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, et al. (2006) Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 43: 1317–1325 DOI:10.1002/hep.21178.
19. Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, et al. (2007) FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology 46: 32–36 DOI:10.1002/hep.21669.

20. Sebastiani G, Vario A, Guido M, Alberti A (2008) Performance of noninvasive markers for liver fibrosis is reduced in chronic hepatitis C with normal transaminases. J Viral Hepat 15: 212–218 DOI:10.1111/j.1365-2893.2007.00932.x.
21. Vallet-Pichard A (2008) Predictive value of FIB-4 versus Fibrotest, APRI, FIBROINDEX and FORNS to noninvasively estimate fibrosis in hepatitis C. Hepatology 47: 763.
22. Nunes D, Fleming C, Offner G, Craven D, Fix O, et al. (2010) Noninvasive markers of liver fibrosis are highly predictive of liver-related death in a cohort of HCV-infected individuals with and without HIV infection. Am J Gastroenterol 105: 1346–1353 DOI:10.1038/ajg.2009.746.
23. Sanvisens A, Fuster D, Serra I, Tor J, Tural C, et al. (2011) Estimated Liver Fibrosis and its Impact on All-Cause Mortality of HCV-Monoinfected and HCV/HIV-Coinfected Drug Users. Curr HIV Res 9:256–262.
24. Saitz R (2005) Clinical practice. Unhealthy alcohol use. N Engl J Med 352: 596–607 DOI:10.1056/NEJMcp042262.
25. National Institute in Alcohol Abuse and Alcoholism (1995) The Physician's guide to helping patients with alcohol problems. *USDHHS NIH Publ* 95–3769.
26. Muga R, Roca J, Egea JM, Tor J, Sirera G, et al. (2000) Mortality of HIV-positive and HIV-negative heroin abusers as a function of duration of injecting drug use. J Acquir Immune Defic Syndr 2000; 23: 332–338.
27. Hutchinson SJ, Bird SM, Goldberg DJ (2005) Influence of alcohol on the progression of hepatitis C virus infection: a meta-analysis. Clin Gastroenterol Hepatol 3: 1150–1159.
28. Siu L, Foont J, Wands JR (2009) Hepatitis C virus and alcohol. Semin Liver Dis 29: 188–199 DOI:10.1055/s-0029-1214374.
29. Tural C, Fuster D, Tor J, Ojanguren I, Sirera G, et al. (2003) Time on antiretroviral therapy is a protective factor for liver fibrosis in HIV and hepatitis C virus (HCV) co-infected patients. J Viral Hepat 10: 118–125.
30. Merchante N, Giron-Gonzalez JA, Gonzalez-Serrano M, Torre-Cisneros J, Garcia-Garcia JA, et al. (2006) Survival and prognostic factors of HIV-infected patients with HCV-related end-stage liver disease. AIDS 20: 49–57.
31. Sulkowski MS, Thomas DL, Chaisson RE, Moore RD (2000) Hepatotoxicity associated with antiretroviral therapy in adults infected with human immunodeficiency virus and the role of hepatitis C or B virus infection. JAMA 283: 74–80.
32. Thomas DL, Astemborski J, Rai RM, Anania FA, Schaeffer M, et al. (2000) The natural history of hepatitis C virus infection: host, viral, and environmental factors. JAMA 284: 450–456.
33. Corey KE, Kane E, Munroe C, Barlow LL, Zheng H, et al (2009) Hepatitis C virus infection and its clearance alter circulating lipids: implications for long-term follow-up. Hepatology 50: 1030–7 DOI:10.1002/hep.23219.
34. Inglesby TV, Rai R, Astemborski J, Gruskin L, Nelson KE, et al (1999) A prospective, community-based evaluation of liver enzymes in individuals with hepatitis C after drug use. Hepatology 29: 590–596 DOI:10.1002/hep.510290219.

Association Between Chromosome 9p21 Variants and the Ankle-Brachial Index Identified by a Meta-Analysis of 21 Genome-Wide Association Studies

Joanne M. Murabito, MD, ScM*; Charles C. White, MPH*; Maryam Kavousi, MD, MSc*; Yan V. Sun, PhD*; Mary F. Feitosa, PhD*; Vijay Nambi, MD*; Claudia Lamina, PhD*; Arne Schillert, PhD*; Stefan Coassin, PhD; Joshua C. Bis, PhD; Linda Broer, MSc; Dana C. Crawford, PhD; Nora Franceschini, MD, MPH; Ruth Frikke-Schmidt, MD, PhD; Margot Haun, MSc; Suzanne Holewijn, PhD; Jennifer E. Huffman, MSc; Shih-Jen Hwang, PhD; Stefan Kiechl, MD; Barbara Kollerits, PhD, MPH; May E. Montasser, PhD; Ilja M. Nolte, PhD; Megan E. Rudock, PhD; Andrea Senft, MSc; Alexander Teumer, PhD; Pim van der Harst, MD, PhD; Veronique Vitart, PhD; Lindsay L. Waite, MS; Andrew R. Wood, MRes; Christina L. Wassel, PhD; Devin M. Absher, PhD; Matthew A. Allison, MD, MPH; Najaf Amin, PhD; Alice Arnold, PhD; Folkert W. Asselbergs, MD, PhD; Yurii Aulchenko, PhD; Stefania Bandinelli, MD; Maja Barbalic, PhD; Mladen Boban, MD, PhD; Kristin Brown-Gentry, MS; David J. Couper, PhD; Michael H. Criqui, MD, MPH; Abbas Dehghan, MD, PhD; Martin den Heijer, MD, PhD; Benjamin Dieplinger, MD; Jingzhong Ding, PhD; Marcus Dörr, MD; Christine Espinola-Klein, MD; Stephan B. Felix, MD; Luigi Ferrucci, MD, PhD; Aaron R. Folsom, MD; Gustav Fraedrich, MD; Quince Gibson, MBA; Robert Goodloe, MS; Grgo Gunjaca, MD; Meinhard Haltmayer, MD; Gerardo Heiss, MD, PhD; Albert Hofman, MD, PhD; Arne Kieback, MD; Lambertus A. Kiemeney, PhD; Ivana Kolcic, MD, PhD; Iftikhar J. Kullo, MD; Stephen B. Kritchevsky, PhD; Karl J. Lackner, MD; Xiaohui Li, MD, MSc; Wolfgang Lieb, MD, MSc; Kurt Lohman, Mstat; Christa Meisinger, MD, MPH; David Melzer, MD, PhD; Emile R. Mohler III, MD; Ivana Mudnic, MD; Thomas Mueller, MD; Gerjan Navis, MD, PhD; Friedrich Oberhollenzer, MD; Jeffrey W. Olin, MD; Jeff O'Connell, PhD; Christopher J. O'Donnell, MD, MPH; Walter Palmas, MD, MS; Brenda W. Penninx, PhD; Astrid Petersmann, MD, PhD; Ozren Polasek, MD, PhD; Bruce M. Psaty, MD, PhD; Barbara Rantner, MD, PhD; Ken Rice, PhD; Fernando Rivadeneira, MD, PhD; Jerome I. Rotter, MD; Adrie Seldenrijk, PhD; Marietta Stadler, MD; Monika Summerer, PhD; Toshiko Tanaka, PhD; Anne Tybjaerg-Hansen, MD, DMSc; Andre G. Uitterlinden, PhD; Wiek H. van Gilst, PhD; Sita H. Vermeulen, PhD; Sarah H. Wild, MB, BChir, PhD; Philipp S. Wild, MD; Johann Willeit, MD; Tanja Zeller, PhD; Tatijana Zemunik, MD, PhD; Lina Zgaga, MD, PhD; Themistocles L. Assimes, MD, PhD; Stefan Blankenberg, MD; Eric Boerwinkle, PhD; Harry Campbell, MD; John P. Cooke, MD, PhD; Jacqueline de Graaf, MD, PhD; David Herrington, MD, MHS; Sharon L.R. Kardia, PhD; Braxton D. Mitchell, PhD; Anna Murray, PhD; Thomas Münzel, MD; Anne B. Newman, MD, MPH; Ben A. Oostra, PhD; Igor Rudan, MD, PhD, MPH; Alan R. Shuldiner, MD; Harold Snieder, PhD; Cornelia M. van Duijn, PhD; Uwe Völker, PhD; Alan F. Wright, PhD; H.-Erich Wichmann, MD, PhD; James F. Wilson, DPhil; Jacqueline C.M. Witteman, PhD; Yongmei Liu, MD, PhD*; Caroline Hayward, PhD*; Ingrid B. Borecki, PhD*; Andreas Ziegler, PhD*; Kari E. North, PhD*; L. Adrienne Cupples, PhD*; Florian Kronenberg, MD*

Background—Genetic determinants of peripheral arterial disease (PAD) remain largely unknown. To identify genetic variants associated with the ankle-brachial index (ABI), a noninvasive measure of PAD, we conducted a meta-analysis of genome-wide association study data from 21 population-based cohorts.

Methods and Results—Continuous ABI and PAD (ABI ≤0.9) phenotypes adjusted for age and sex were examined. Each study conducted genotyping and imputed data to the ≈2.5 million single nucleotide polymorphisms (SNPs) in HapMap. Linear and logistic regression models were used to test each SNP for association with ABI and PAD using additive genetic models. Study-specific data were combined using fixed effects inverse variance weighted meta-analyses. There were a total of 41 692 participants of European ancestry (≈60% women, mean ABI 1.02 to 1.19), including 3409 participants with PAD and with genome-wide association study data available. In the discovery meta-analysis, rs10757269 on chromosome 9 near *CDKN2B* had the strongest association with ABI ($\beta=-0.006$, $P=2.46\times10^{-8}$). We sought replication of the 6 strongest SNP associations in 5 population-based studies and 3 clinical samples (n=16 717). The association for rs10757269 strengthened in the combined discovery and replication analysis ($P=2.65\times10^{-9}$). No other SNP associations for ABI or PAD achieved genome-wide significance. However, 2 previously reported candidate genes for PAD and 1 SNP associated with coronary artery disease were associated with ABI: *DAB21P* (rs13290547, $P=3.6\times10^{-5}$), *CYBA* (rs3794624, $P=6.3\times10^{-5}$), and rs1122608 (*LDLR*, $P=0.0026$).

Conclusions—Genome-wide association studies in more than 40 000 individuals identified 1 genome wide significant association on chromosome 9p21 with ABI. Two candidate genes for PAD and 1 SNP for coronary artery disease are associated with ABI. **(*Circ Cardiovasc Genet*. 2012;5:100-112.)**

Key Words: cohort study ■ genetic association ■ genome-wide association study ■ meta-analysis ■ peripheral vascular disease

Received August 1, 2011; accepted December 9, 2011.

A list of the institutions and affiliations for the authors of this report may be found in the Appendix at the end of this article.

Guest Editor for this article was Barbara V. Howard, PhD.

*These authors contributed equally.

The online-only Data Supplement is available at http://circgenetics.ahajournals.org/lookup/suppl/doi:10.1161/CIRCGENETICS.111.961292/-/DC1.

Correspondence to Dr. Joanne M. Murabito, Framingham Heart Study, 73 Mount Wayte Ave, Suite 2, Framingham, MA 01701 (E-mail murabito@bu.edu); or Dr. Florian Kronenberg, Division of Genetic Epidemiology, Innsbruck Medical University, Schöpfstr. 41, 6020 Innsbruck, Austria (E-mail Florian.Kronenberg@i-med.ac.at).

Peripheral arterial disease (PAD) affects approximately 27 million people in Europe and North America,[1] and is associated with increased risk for myocardial infarction, stroke, and mortality.[2–6] Measurement of ankle and arm blood pressures with a Doppler device and calculation of the ankle-brachial index (ABI) is a simple and reliable method to detect PAD. An ABI ≤0.90 is indicative of definite PAD.[7] In previous work, the Ankle-Brachial Index Collaboration demonstrated a reverse J-shaped relationship of ABI with mortality and coronary events, with a low risk ABI ranging from 1.11 to 1.40.[8]

Clinical Perspective on p 112

Little is known about genetic susceptibility to PAD, but familial aggregation and heritability estimates suggest a significant genetic component.[9–13] A study of 112 biological candidate genes identified only 2 single nucleotide polymorphisms (SNPs) in *NOS3* significantly associated with ABI.[14] The candidate gene approach to identify novel genetic variants for PAD has been limited by modest study sample size, relatively small number of genes examined, and lack of replication in independent samples.[13]

Genome-wide association studies (GWAS) have led successfully to the discovery of novel genetic variants for several common diseases, including coronary artery disease (CAD).[15] The association between genetic variants on chromosome 9p21 and CAD has demonstrated replication,[16,17] persistent association across race or ethnicity,[18] and association with other vascular diseases.[19–21] Notably, GWAS of subclinical atherosclerosis phenotypes, such as intima-medial thickness or ABI, are sparse. Therefore, we conducted a meta-analysis of GWAS findings for ABI within an international consortium of 21 population-based cohort studies that included 41 692 participants of European ancestry, among whom 3409 participants had PAD (ABI ≤0.90). We conducted replication analyses of our strongest findings in over 16 000 individuals from population-based cohort studies and clinically based samples of PAD. We hypothesized that this approach would lead to the unbiased identification of genetic variants associated with ABI. Further, we hypothesized that some genetic variants for ABI would be identical to those reported to be associated with CAD or its risk factors given shared underlying biological pathways, while some genetic variants would be associated uniquely with PAD.

Methods

Discovery Studies

Our analyses were conducted within the international Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium,[22] and included 4 of the 5 original CHARGE cohorts: Atherosclerosis Risk in Communities Study (ARIC, n=7630), the Cardiovascular Health Study (CHS, n=3193), the Framingham Heart Study (FHS, n=3572), and the Rotterdam Study (RS-I, n=5169 and RS-II, n=1642). Ten additional population-based cohorts joined the collaboration for analysis of ABI phenotypes: the Family Heart Study (FamHS, n=1736), Genetic Epidemiology Network of Arteriopathy Study (GENOA, n=991), Gutenberg Heart Study (GHS, n=3122), Health, Aging, and Body Composition (Health ABC, n=1564), the Invecchiare in Chianti Study (InCHIANTI, n=1130), Cooperative Health Research in the Region of Augsburg (KORA F3, n=1581 and KORA F4, n=1407), Netherlands Study of Anxiety and Depression (NESDA, n=1612), Nijmegen Biomedical Study (NBS, n=544), and the Study of Health in Pomerania (SHIP, n=543). A further 6 studies derived from population isolates also were available for the analyses: Amish Study (Amish, n=1183), Croatia-Vis (n=897), Croatia-Korcula (n=851), Croatia-Split (n=499), Erasmus Rucphen Family Study (ERF, n=2133), and the Orkney Complex Disease Study (ORCADES, n=693). For all studies participating in the meta-analyses, each participant self-identified as European or European-American and provided written informed consent, and the Institutional Review Board at the parent institution for each respective cohort approved the study protocols. More detailed study-specific information is provided in the online-only Data Supplement Methods.

Ankle-Brachial Index Phenotypes

Ankle and brachial blood pressure measurements for each participating study were obtained from the baseline examination or the first examination in which the measurement was obtained. Details on the ABI protocol used and the calculation performed in each study are provided in online-only Data Supplement Table I. To calculate the ABI for each leg, the systolic blood pressure at each ankle was divided by the systolic blood pressure in the arm. If the systolic blood pressure was measured in both arms, the higher arm reading was used in the ABI calculation. If replicate readings were obtained, the mean of the 2 measurements for each limb was used to calculate the ABI, with the exception of InCHIANTI, which used the higher of the 2 readings of each measurement set to calculate the ABI. The lower of the ABIs from the 2 legs was used for analysis. In ARIC and FamHS, the ABI was measured in only 1 leg, chosen at random. Participants with an ABI >1.40 were excluded because this high ABI may represent medial sclerosis, fibrocalcific disease secondary to diabetes mellitus, or other causes of noncompressible vessels.

To maximize the sample size and the power to detect genetic variants with modest effects, and to examine the entire range of ABI values given the recent evidence of increased cardiovascular disease risk associated with ABI values up to 1.1,[8] we examined the continuous range of ABI <1.40. As a secondary analysis to provide a clinical phenotype, we defined PAD as ABI ≤0.90 and conducted a case (ABI ≤0.9)/control; ABI >0.90 and <1.40) comparison analysis.

Genotyping and Imputation

Different genotyping platforms were used by the 21 studies (online-only Data Supplement Table II). Each study imputed the genotype "dosage" (0 to 2) for the expected number of alleles for ≈2.5 million Phase II HapMap CEU SNPs for each participant using currently available imputation methods.[23] CHS used BIMBAM (available at http://stephenslab.uchicago.edu/software.html),[24] GHS, InCHIANTI, NESDA, and SHIP used IMPUTE,[25] and all other cohorts used MACH (http://www.sph.umich.edu/csg/abecasis/MaCH/).

Statistical Analysis

We devised a GWAS analysis plan for the ABI and PAD phenotypes that each study independently implemented. Sex-specific and age-adjusted residuals of ABI were created from linear regression models and used as phenotypes in the analysis. No transformation of the ABI measure was performed before analysis. In FHS, residuals also were obtained separately in the original and offspring cohorts. Multi-site studies (ARIC, CHS, and FamHS) additionally adjusted for field study site. Each SNP was tested for association with ABI in additive genetic models using linear regression. The Amish Study, FamHS, FHS, and GENOA cohorts used linear mixed effects models to account for familial correlations. Croatia-Vis, Croatia-Korcula, Croatia-Split, ERF, and ORCADES used the "mmscore" function of the GenABEL package for R statistical software for the association test under an additive model. This score test for a family-based association takes into account pedigree structure and allows unbiased estimations of SNP allelic effect when relatedness is present between examinees. Logistic regression adjusting for age and sex was used to test each SNP for association with the PAD phenotype. The FamHS,

FHS, and GENOA cohorts used generalized estimating equations clustering on family to account for family correlations.

A genome-wide meta-analysis using a fixed effects approach with inverse variance weighting was then conducted in METAL[26] [www.sph.umich.edu/csg/abecasis/metal] for 2 669 158 SNPs in the meta-analysis, excluding the population isolates (2 670 732 SNPs including the population isolates) that met imputation and quality control criteria (online-only Data Supplement Table II). Before meta-analysis, genomic control was applied to each study. The association of ABI per each additional risk allele was quantified by the regression slope (β), its standard error [SE(β)], and the corresponding probability value. We calculated a meta-analysis odds ratio for each of the most significant SNP associations for PAD. The meta-analysis odds ratio estimates the increase in odds of PAD for each additional copy of the risk allele of the SNP. SNP associations were considered to be significant on a genome-wide level at $P<5\times10^{8}$.[27,28] Standardized gene and SNP annotations were created using a PERL script.[29] We also tested for heterogeneity of study specific regression parameters using Cochran Q statistic. Because of concerns about heterogeneity, we conducted analyses of nonisolate studies and of the full group of studies. We selected SNPs for replication using results from the meta-analysis, excluding the population isolates, because the available replication samples did not include isolates. We excluded SNP association results if the total meta-analysis sample was less than 20 000 and if the average minor allele frequency of the SNP was <5%.

Replication

We sought to replicate independent SNP associations for ABI that attained genome-wide significance (1 region), SNPs with suggestive associations (5 regions, $P<10^{-5}$), and bioinformatics data supporting the signal. The bioinformatic analyses are described in detail in the online-only Data Supplement Material. In addition, we sought to replicate 1 SNP associated with both ABI and PAD at $P<10^{-4}$. The replication studies included 5 population-based studies and 3 clinically-based studies, including a total of over 16 000 participants: the Bruneck Study (n=786), the Copenhagen City Heart Study (CCHS, n=5330), the Multi-Ethnic Study of Atherosclerosis (MESA, n=2611), the National Health and Nutrition Examination Surveys (NHANES 1999–2002, n=2335), Prevention of Renal and Vascular End-stage disease (PREVEND, n=3691) cohort, Cardiovascular Disease in Intermittent Claudication (CAVASIC, n=443) Study, Genetic Determinants of Peripheral Arterial Disease (GenePAD, n=850), and the Linz Peripheral Arterial Disease (LIPAD, n=671) Study. Each collaborating study was provided with a SNP list and a detailed analysis plan. MESA and PREVEND used *in silico* genotyping (online-only Data Supplement Table II), and the remaining studies genotyped the SNPs using Taqman assays or Sequenom. Relative excess heterozygosity analysis demonstrated that all genotyped SNPs were compatible with Hardy-Weinberg equilibrium at the nominal 5% test-level (online-only Data Supplement Table III).[30]

Examination of Candidate Genes Associated With Peripheral Artery Disease and Coronary Artery Disease/Myocardial Infarction

We selected candidate genes for ABI or PAD from the published literature using PubMed search terms "([ankle-brachial index] OR [peripheral arterial disease]) AND polymorphism." Association studies with at least 100 cases and 100 controls were included regardless of whether the original study results were positive or negative. Using the discovery meta-analysis results for ABI, we then identified the most strongly associated SNPs based on probability values within the gene region ±100 kb upstream or downstream of the candidate gene. Because of the high correlation of imputed genotypes, the effective number of loci were calculated for each gene region[31] using the genotype scores from the KORA F4 Study (online-only Data Supplement Methods). Bonferroni correction of probability values then was applied in each region using the effective number of loci. Subsequently, false discovery rates (FDR) were calculated using these corrected probability values, accounting for the number of gene regions examined (online-only Data Supplement Methods). Lastly, we examined the association with ABI of 30 SNPs strongly associated with CAD in recent GWAS.[32–34] Our ABI discovery meta-analysis did not include 2 of the 30 SNPs (rs17465637 and rs3798220), and we were unable to identify proxy SNPs available in our data. Using the probability values for the 28 SNPs in our discovery meta-analysis, we then calculated the FDR for each CAD SNP, accounting for the 28 regions examined.

Results

Study Sample

The study sample included 41 692 participants of European ancestry (56% women, 6256 from population isolates) with ABI data and genome-wide genotyping. Participant characteristics at the time of ABI measurement for each cohort are provided in online-only Data Supplement Table IV. Across the studies the mean age ranged from 41.8 years to 73.8 years, the mean ABI ranged from 1.02 to 1.19, and 8.2% (n=3409) had PAD (ABI <0.9). Characteristics of the replication samples were similar to the discovery set (online-only Data Supplement Table V).

ABI-SNP Associations

We conducted a meta-analysis with (n=41 692) and without (n=35 434) the population isolates (online-only Data Supplement Figures I and II, QQ-plots and Manhattan plots, and study-specific lambdas ranged from 0.997 to 1.044). Our primary meta-analysis excluded studies from population isolates because of concern for study heterogeneity and the lack of availability of replication samples from isolates. The strongest SNP association for ABI was rs10757269 on chromosome 9 near *CDKN2B* ($\beta=-0.006$, $P=2.46\times10^{-8}$, P for heterogeneity=0.23, Table 1; meta-analysis results, including the population isolates, online-only Data Supplement Table VII). Among the 96 SNP associations for ABI with $P<10^{-5}$, 79 were located in the chromosome 9p21 region (online-only Data Supplement Table VI). The ABI SNP rs10757269 is in strong linkage disequilibrium (LD), with several SNPs in the region previously reported to be associated with CAD or myocardial infarction ($r^2>0.8$), but this ABI SNP is not in LD with SNPs previously associated with the type 2 diabetes mellitus (Figure 1). We repeated the meta-analysis to examine the association between ABI and rs10757269, first adjusting for CAD and then excluding individuals with CAD among the nonisolate studies. The association remained but was no longer genome-wide significant (adjusting for CAD: $P=5.56\times10^{-6}$; excluding CAD: $P=3.79\times10^{-5}$). Next, we sought to replicate the association between rs10757269 and ABI in both population-based and clinically-based samples (n=16 717). The magnitude and direction of the association in the replication studies was similar to the discovery set ($\beta=-0.0035$, $P=0.0176$), providing evidence of replication. In the combined stage 2 discovery plus replication meta-analysis, the ABI-rs10757269 association became stronger ($P=2.65\times10^{-9}$). The study-specific estimates of effect for the discovery studies, population isolates, replication studies, and overall discovery plus replication meta-analyses are presented in Figure 2. Two studies among the population isolates (the Amish Study and Croatia-Split) had effect estimates in the

Table 1. Meta-Analysis Results: ABI-SNP Associations with $P<10^{-5}$ in the Primary Discovery Analysis With Population Isolates Excluded

SNP	Chr	Physical Position	Closest Gene	Risk/Non-Risk Allele	Risk Allele Frequency	Meta-Analysis	N	Beta	SE	*P* Value	P_{het}
rs10757269	9	22062264	*CDKN2B*	G/A	0.49	ABI discovery	35 036	−0.0056	0.001	2.46E-08	0.23
						ABI replication	16 672	−0.0035	0.0015	1.76E-02	0.67
						ABI combined	51 708	−0.0049	0.0008	2.65E-09	0.38
						PAD† discovery	34 555	0.0849	0.0296	4.15E-03	0.32
rs4659996	1	238912747	*GREM2*	A/G	0.48	ABI discovery	28 087	−0.006	0.0012	4.44E-07	0.34
						ABI replication	16 658	−0.0018	0.0016	2.67E-01	0.65
						ABI combined	44 745	−0.0045	0.001	2.12E-06	0.32
						PAD discovery	27 574	0.0725	0.0319	2.31E-02	0.52
rs7003385	8	41705907	*ANK1*‡	T/C	0.67	ABI discovery	35 375	−0.0053	0.0011	5.24E-07	0.49
						ABI replication	16 690	−0.002	0.0016	2.20E-01	0.52
						ABI combined	52 065	−0.0043	0.0009	1.11E-06	0.43
						PAD discovery	34 903	0.0838	0.0314	7.57E-03	0.24
rs819750	1	99469651	*LPPR4*‡	G/T	0.12	ABI discovery	35 278	−0.007	0.0015	3.64E-06	0.51
						ABI replication	16 660	0.0022	0.0023	3.22E-01	0.99
						ABI combined	51 938	−0.0041	0.0013	1.01E-03	0.31
						PAD discovery	34 780	0.1068	0.0437	1.45E-02	0.06
rs9485528	6	102221473	*GRIK2*‡	A/G	0.17	ABI discovery	35 339	−0.0061	0.0013	4.63E-06	0.78
						ABI replication	16 679	0.0002	0.002	9.24E-01	0.63
						ABI combined	52 018	−0.0041	0.0011	1.77E-04	0.48
						PAD discovery	34 850	0.1172	0.0380	2.02E-03	0.80
rs722453	7	84037497	*SEMA3A*	G/A	0.42	ABI discovery	26 200	−0.0054	0.0012	6.43E-06	0.69
						ABI replication	6300	−0.0046	0.0025	5.74E-02	0.08
						ABI combined	32 500	−0.0052	0.0011	1.02E-06	0.59
						PAD discovery	25 706	0.0575	0.0318	7.05E-02	0.63
rs16824978	2	211380306	*CPS1*	T/C	0.25	ABI discovery	34 950	−0.0054	0.0012	7.77E-06	0.37
						ABI replication	14 340	0.0000	0.0019	9.94E-01	0.22
						ABI combined	49 290	−0.0039	0.001	1.48E-04	0.11
						PAD discovery	34 518	0.0760	0.0343	2.65E-02	0.39

P_{het} indicates *P* value for heterogeneity; ‡, SNP is located within the gene; rs819750 is within 60kb of the gene; †, PAD discovery: ABI <0.9 vs ABI >0.9. Chr indicates chromosome.

opposite direction to the other studies. None of the other SNP associations for ABI achieved genome-wide significance. The significance of the associations for the additional SNPs chosen for replication diminished in the discovery plus replication meta-analysis (Table 1, online-only Data Supplement Table VII).

PAD-SNP Associations

None of the SNP associations for the PAD phenotype (defined by an ABI ≤0.9) achieved genome-wide significance (Table 2; for meta-analysis results including population isolates see online-only Data Supplement Table VIII). The strongest association was found for rs6584389 on chromosome 10 near the *PAX2* gene (odds ratio 1.17, 95% confidence interval 1.10, 1.25, $P=2.34\times10^{-6}$). Of note, the chromosome 9 SNP rs10757269 association with PAD was in a direction consistent with the ABI association but did not achieve statistical significance (Table 1, $\beta=0.0849$, $P=0.004$, increasing the odds of PAD).

Overlap in SNP Associations for ABI and PAD

While the directions of effect for the ABI SNPs in Table 1 were consistent with the PAD association result (lower ABI, increased odds of PAD), there was little overlap in the top associations for the 2 phenotypes. Only 3 regions marked by SNPs in or near *IDE* (10q23–q25), *DAB21P* (9q33.2), and *GRAMD1C* (3q13.31), in addition to the chromosome 9p21 region, showed association with both ABI and PAD at the $P<10^{-4}$ level (online-only Data Supplement Table IX). SNP rs7100623 in *IDE* demonstrated the strongest novel association with both ABI ($\beta=-0.005$, $P=1.89\times10^{-5}$) and PAD ($\beta=0.139$, $P=8.39\times10^{-5}$) at $P<10^{-4}$; however, the association probability value was not significant in the replication stage, and diminished in the combined discovery plus replication meta-analysis.

Examination of PAD Candidate Genes

Among the 55 candidate genes or regions previously tested for association with ABI or PAD, 8 regions showed nominally significant probability values ($P<0.05$) after correction

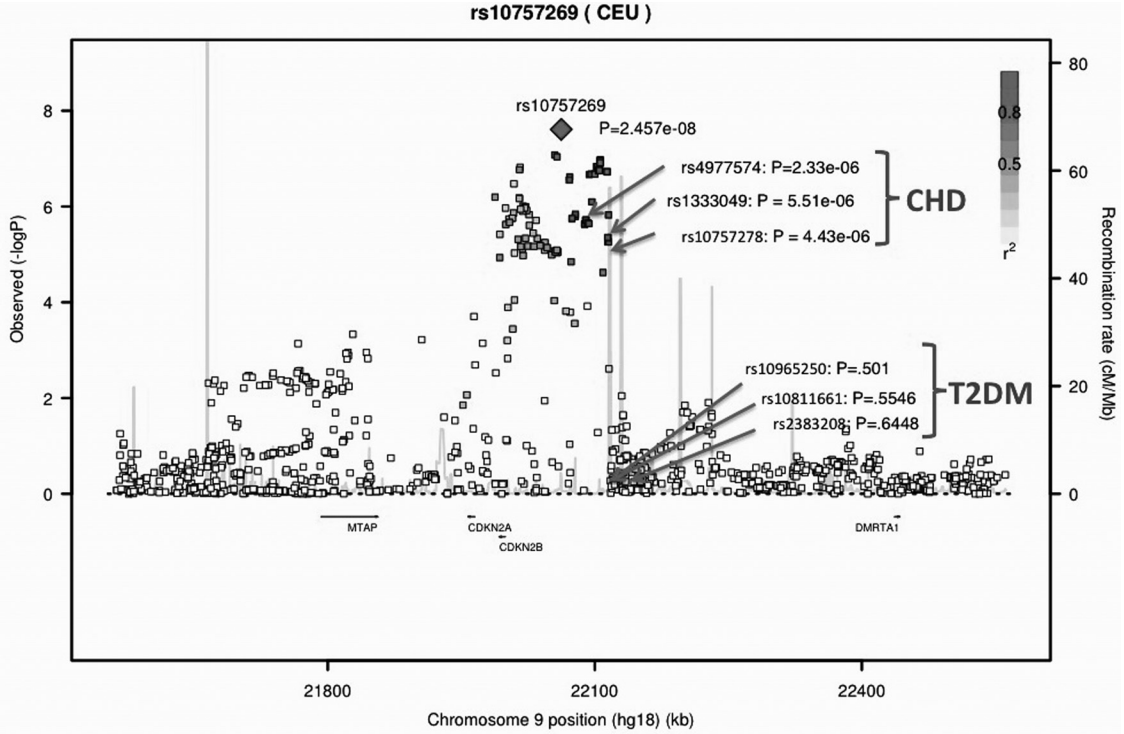


Figure 1. Genomic context of the genome-wide significant signal at chromosome 9p21 plotted against the $-\log_{10} P$ values. r^2 is between the top signal (rs10757269) and each SNP is shown. SNPs previously reported from genome-wide association studies (GWAS) to be associated with coronary artery disease (CHD, arrows), type 2 diabetes (T2DM, arrows), and *P* value for association with ankle-brachial index are shown. Chromosome positions are based on build hg18.

for the number of effective loci for each gene region. After accounting for the number of regions examined using a false discovery rate (FDR <0.10), we found evidence of association between ABI and *CYBA* (rs3794624, uncorrected $P=6.3\times10^{-5}$, corrected $P=0.0036$, FDR=0.0665) and *DAB21P* (rs13290547, uncorrected $P=3.6\times10^{-5}$, corrected $P=0.0035$, FDR=0.0665), in addition to the chromosome 9p21 locus (rs1333049) reported to be associated with ABI (Table 3).[35] We found no evidence of association between ABI and any of the other candidate genes previously tested for association with ABI or PAD (online-only Data Supplement Table X).

Examination of Coronary Artery Disease/ Myocardial Infarction Candidate Genes

Among the 30 SNPs previously reported by GWAS to be associated with CAD or myocardial infarction, 28 SNPs were available in our discovery meta-analysis of ABI, and 2 of these SNPs demonstrated an association (FDR <0.10) with ABI, including rs4977574 near *CDKN2B* ($P=2.33\times10^{-6}$) and rs1122608 in *LDLR* ($P=0.0026$) (Table 3, online-only Data Supplement Table XI).

Discussion

Our GWAS meta-analysis for ABI conducted in more than 40 000 adults of European ancestry has several notable findings. First, we identified and replicated 1 genome-wide significant association between a SNP in the chromosome 9p21 region and ABI. No other ABI-SNP associations achieved genome-wide significance. Second, in our discovery sample, over 3000 adults had PAD (ABI ≤0.9); however, none of the SNP associations were significant. Third, the directions of effect were consistent across the 2 phenotypes for the most significant ABI SNPs (lower ABI, increased odds of PAD): however, we observed minimal overlap in the top SNP associations for ABI and PAD. Finally, the effect size for the 9p21 SNP was modest. The association itself is, however, intriguing, and may provide insights into the biological mechanisms contributing to generalized atherosclerosis.

Chromosome 9p21 Locus and Atherosclerosis Susceptibility

Common genetic variants in the 9p21 locus are associated strongly with myocardial infarction and CAD,[17,33,36] and confer risk for other atherosclerotic diseases including stroke,[19] cerebral and abdominal aortic aneurysm,[20,21] and clinically diagnosed PAD; however, the relation with PAD was diminished when coronary artery disease cases were excluded.[20] SNP associations at the 9p21 locus with subclinical measures of atherosclerosis have been conflicting. Initially, no association was observed with carotid intima-medial thickness or flow mediated dilation in young or older adults;[37,38] however, more recent reports demonstrate an association with the development and progression of carotid atherosclerosis[39] and with the suggestion of a stronger effect in men.[40] To further investigate the ABI-9p21 SNP association noted in this study, we conducted the meta-analysis after adjusting for CAD and after exclusion of individuals with CAD. Not surprisingly, the association persisted but was no

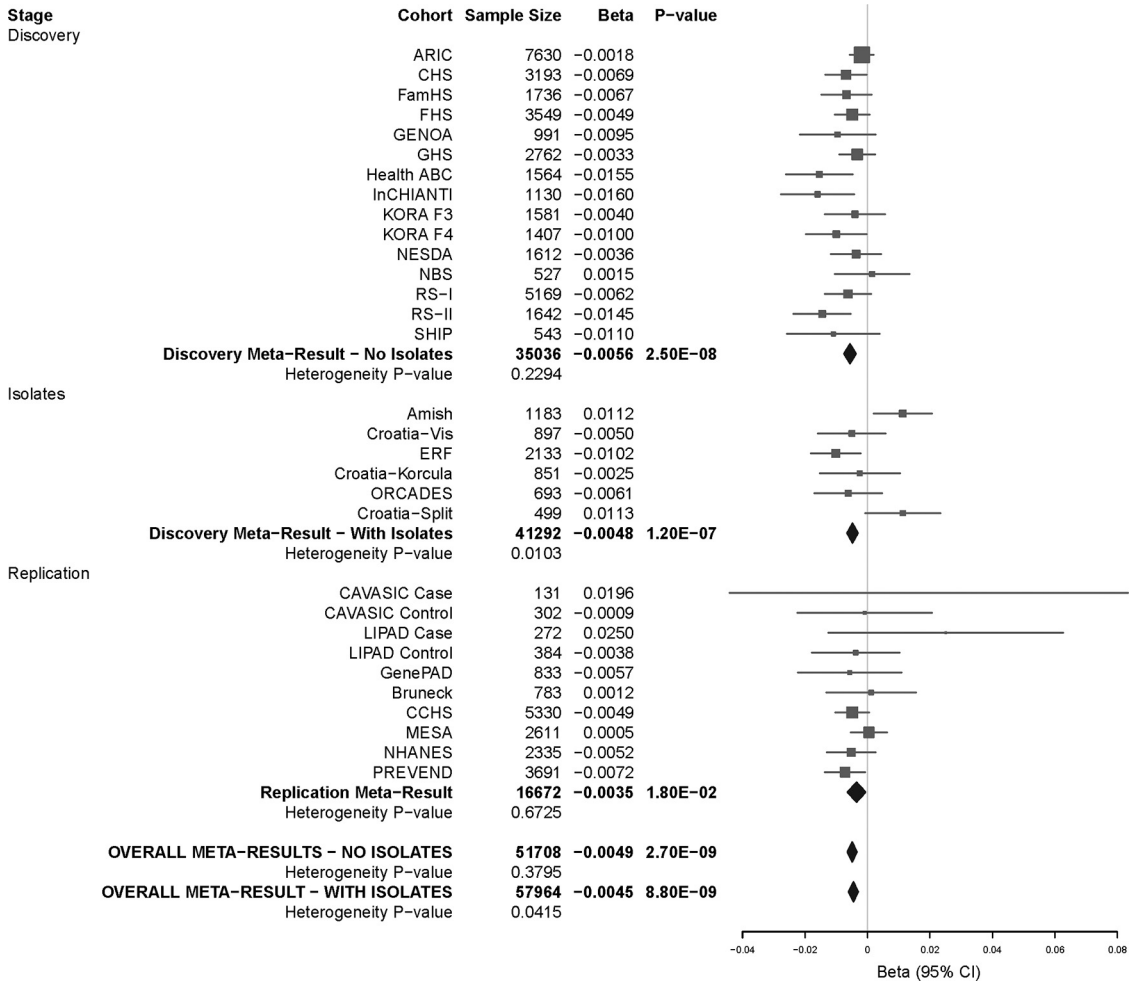


Figure 2. Ankle-brachial index-chromosome 9p21 (rs10757269) association: study-specific estimates of effect for the discovery studies, population isolates, replication studies, and overall discovery and replication meta-analyses.

longer genome-wide significant. Both CAD and PAD are manifestations of underlying atherosclerosis, and nearly two thirds of individuals with PAD have coexisting coronary or cerebrovascular disease.[41] One previous report conducted in 3 studies of older adults identified an association between a variant at 9p21 and lower ABI, as well as an increased risk for PAD.[35] The primary effect of the chromosome 9p21 region may be on the atherosclerotic process itself, and there are likely to be many other factors, both genetic and environmental, that determine whether it manifests as CAD, PAD, or another clinical atherosclerotic phenotype. The primary biological mechanism underlying the association with ABI is unknown but appears to be independent of 2 major PAD risk factors, diabetes and smoking, as the ABI SNP in the 9p21 region we identified is not in linkage disequilibrium with the SNPs in the region associated with diabetes risk[42,43] or smoking-related behaviors.[44] The mechanism may be related to modulation of platelet reactivity,[45] atheroma formation, plaque instability, thrombosis, or biological processes not yet identified.[46] The SNP associated with ABI is nearest to

Table 2. Meta-Analysis Results: SNP Associations for PAD (ABI ≤0.9 vs ABI >0.9) With $P<10^{-5}$ With Population Isolates Excluded

SNP	Chr	Physical Position	Closest Gene	Risk/Non-Risk Allele	Risk Allele Frequency	N	OR	95% Confidence Interval	*P* Value	P_{het}
rs6584389	10	102459392	*PAX2*	C/A	0.50	24 474	1.17	(1.10, 1.25)	2.34E-06	0.37
rs9998941	4	162544312	*FSTL5**	A/G	0.23	34 670	1.18	(1.10, 1.27)	2.34E-06	0.61
rs11751656	6	42751046	*UBR2**	G/A	0.07	27 470	1.61	(1.32, 1.96)	2.46E-06	0.75
rs4535726	8	68938371	*DEPDC2*	T/C	0.20	34 915	1.18	(1.10, 1.26)	3.79E-06	0.01
rs2090205	17	73897869	*PGS1**	A/C	0.24	34 912	1.16	(1.09, 1.24)	5.01E-06	0.17
rs11933540	4	25729099	*RBPJ*	C/T	0.30	34 830	1.15	(1.08, 1.23)	9.86E-06	0.08

P_{het} indicates *P* value for heterogeneity.
*SNP is located within the gene. Chr indicates chromosome.

Table 3. Literature-Reported Candidate Genes for Peripheral Artery Disease and Coronary Artery Disease and Their Association With Ankle-Brachial Index in the CHARGE GWAS Discovery Sample (Population Isolates Excluded) With FDR <0.10†

SNP	Chr	Physical Position	Closest Gene	Risk/ Non-Risk Allele	Risk Allele Frequency	N	Beta	SE	*P* Value*	# of effective loci‡	*P* Value Corrected‡	False Discovery Rate‡
PAD genes												
rs10757269	9	22 062 264	*CDKN2B*	G/A	0.51	35036	−0.006	0.001	2.50E-08	69	1.70E-06	9.32E-05
rs3794624	16	87 244 575	*CYBA*	G/A	0.34	31035	−0.005	0.001	6.30E-05	58	3.60E-03	0.0665
rs13290547	9	123 527 316	*DAB2IP*	T/C	0.06	32135	−0.009	0.002	3.60E-05	97	3.50E-03	0.0665
CAD genes												
rs4977574	9	22 088 574	*CDKN2B*	G/A	0.49	35411	−0.0047	0.001	2.33E-06	. . .	. . .	6.52E-05
rs1122608	19	11 024 601	*LDLR*	G/T	0.74	35384	−0.0035	0.001	2.56E-03	. . .	. . .	0.036

**P* value from Discovery GWAS of ABI. Chr indicates chromosome.

†Candidate genes for PAD were selected for testing with ABI if an association study with at least 100 cases and 100 controls was available in the literature, independent of whether the study was positive or negative. Genes for CAD were considered only for testing with ABI if they were identified by recent GWAS to be genome-wide significantly associated with CAD. The table shows only the genes which showed an experiment-wise significant association with ABI after correction for multiple testing. The entire list of genes can be seen in online-only Data Supplement Table X and XI for PAD and CAD genes, respectively.

‡Due to the high correlation of imputed genotype scores, the effective number of loci was calculated for each PAD gene region (31) using the genotype scores from the KORA F4 Study. Bonferroni correction of *P* values then was applied in each region using this number. Furthermore, the corrected *P* value thresholds of significance for 28 CAD loci (tested in online-only Data Supplement Table XI, $\alpha=0.05/28$, 1.85×10^{-3}) and 55 PAD loci (tested in online-only Data Supplement Table X, $\alpha=0.05$/effective number of loci) were calculated. We also calculated a false discovery rate (FDR) using the corrected *P* values accounting for the number of gene regions examined. An FDR <0.10 defined evidence of a significant association.

CDKN2B, a well recognized tumor-suppressor gene that encodes a cyclin-dependent kinase inhibitor and is involved in regulation of the cell cycle. *CDKN2B* is abundantly expressed in human atherosclerotic lesions,[47] and animal models suggest that altered *CDKN2A/B* expression results in abnormal regulation of vascular cell proliferation.[48] Functional studies reveal a long noncoding RNA at this locus named ANRIL, and a mouse model has confirmed the essential role of ANRIL in regulation of *CDKN2B* expression through a cis-acting mechanism.[49,50] ANRIL is implicated in proliferation and senescence.

PAD Candidate Genes

We performed a literature search to identify all candidate gene regions previously investigated for association with PAD or ABI, irrespective of whether the association was reported to be positive or negative. This approach revealed 2 further associated gene regions: *DAB2IP* and *CYBA*. *DAB2IP* rs13290547 was not only associated with ABI, but also with PAD ($P=3.62\times10^{-5}$ and 2.2×10^{-5}, respectively; online-only Data Supplement Table X). The *DAB2IP* gene encodes an inhibitor that is involved in the regulation of cell survival and proliferation. One variant in the *DAB2IP* gene (rs70254486) recently has been detected in a GWAS of abdominal aortic aneurysm.[51] That study also detected an association with PAD as a secondary end point in 3690 cases versus 12 271 controls ($P=3.9\text{x}10^{-5}$). The same SNP showed an association with CVD within a meta-analysis of case control studies.[52] The *CYBA* gene is involved in NADPH oxidase regulation, which contributes to oxidative stress and plays a key role in the pathophysiology of coronary disease. Only 1 report investigated a SNP (rs4673) in this gene for association with PAD among 324 cases and 295 controls, but did not find an association.[53] Our study found an association of rs3794624 ($r^2=0.5$ with rs4673) with continuous ABI, which may indicate that the earlier study likely lacked power to find this association. None of the other gene regions had sufficient evidence for association with continuous ABI in our meta-analysis. Another very wide-reaching approach designed to systematically examine a large number of genes related to intermediate phenotypes of atherosclerosis, such as blood pressure regulation, lipoprotein metabolism, inflammation, oxidative stress, vascular wall biology, obesity, and diabetes, found only eNOS to be significantly associated with ABI.[14] This gene could not be confirmed by our candidate gene examination.

Coronary Candidate Genes

Besides the chromosome 9 locus, 1 other SNP reported to be associated with coronary disease in recent GWAS also showed an association with ABI in our study; rs1122608 in *LDLR*. The *LDLR* gene plays an important role in cholesterol homeostasis, and mutations at this gene have been shown to influence LDL cholesterol levels and the subsequent risk for coronary disease.[54] The association of *LDLR* gene with ABI in our study is a confirmation of the shared biological pathways underlying both subclinical and clinically apparent disease.

Strengths/Limitations

Our meta-analysis represents the largest collaborative effort to date to identify genome-wide SNP associations for variation in ABI and PAD (ABI ≤0.90), and our findings suggest the absence of common variants with large effects on ABI. Use of ABI as our primary phenotype has major advantages of detecting asymptomatic PAD, as the ABI is an objective measurement, whereas clinical PAD requires subjective symptoms of exertional leg discomfort and mobility of the individual. However, several limitations of our meta-analysis merit comment. The blood pressure measurement protocol and ABI calculation was heterogeneous across participating studies. While protocols were standardized within each study,

the studies were not designed to be fully standardized and comparable across studies (online-only Data Supplement Table I). This phenotype heterogeneity may have impacted our ability to detect associations. Furthermore, for many studies, information about a previous revascularization intervention was not available. This lack of data may have resulted in the misclassification of some of the most affected persons by placing them into an ABI range of unaffected individuals and consequently reducing our power to detect true associations. Our sample was restricted to individuals of European ancestry, and thus our findings cannot yet be generalized to individuals of other race or ethnic groups. Furthermore, some PAD susceptibility variants may be race or ethnic specific and only can be uncovered through the study of non-Europeans. For example, African-Americans have a higher prevalence of PAD that cannot be attributed to traditional or novel risk factors.[55] This observation raises the hypothesis that polymorphisms unique to African-Americans partially may be responsible for the higher prevalence of PAD.[55] We did not evaluate gene by environment interactions, which may be especially relevant for cigarette smoking, a strong risk factor for PAD,[56] and a factor known to interact with other genes to modulate atherosclerosis.[57]

Conclusions

In conclusion, a common variant near the *CDKN2B* gene in the chromosome 9p21 locus is associated with a lower ABI. PAD represents a diffuse form of atherosclerosis associated with increased risk for death and incident CVD events. Thus, the identification of genetic variants associated with ABI may provide an important opportunity not only to unravel the biological basis of PAD, but also to improve our understanding of the causes of the variation in degree of atherosclerosis from 1 arterial bed to another. Additional studies are warranted to identify the causal variants in the 9p21 locus and to characterize their functional significance. The search for genes influencing predilection to PAD remains elusive, and alternative approaches are warranted.

Appendix

The following is a list of institutional and study affiliations:

From the NHLBI's Framingham Heart Study, Framingham, MA (J.M.M., S.-J.H., C.J.O., L.A.C.); Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA (J.M.M.); Department of Biostatistics, Boston University, Boston, MA (C.C.W., L.A.C.); Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands (M.K., L.B., N.A., Y.A., A.D., A.H., F.R., A.G.U., C.M.D., J.C.M.W.); Netherlands Genomics Initiative (NGI)-Sponsored Netherlands Consortium for Healthy Aging (NCHA) and Center for Medical Systems Biology, Rotterdam, the Netherlands (M.K., L.B., A.D., A.H., F.R., A.G.U, C.M.D., J.C.M.W.); Department of Epidemiology, Emory University School of Public Health, Atlanta, GA (Y.V.S.); Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO (M.F.F., I.B.B.); Department of Atherosclerosis and Vascular Medicine, Baylor College of Medicine, Houston, TX (V.N.); Genetic Epidemiology, Department of Medicine Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria (C.L., S.C., M.H., B.K., B.R., M.S., F.K.); Institut für Med Biometrie and Statistik, Univ zu Lübeck, Universitätsklinikum Schleswig-Holstein, Lübeck, Germany (A.S., A.S., A.Z.); Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA (J.C.B.); Department of Molecular Physiology and Biophysics, The Center for Human Genetics Research, Vanderbilt University, Nashville, TN (D.C.C.); Department of Epidemiology, University of North Carolina Gillings School of Global Public Health, The University of North Carolina, Chapel Hill, NC (N.F., G.H., K.E.N.); Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark (R.F.-S., A.T-H.); Department of General Internal Medicine, Vascular Medicine, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands (S.H., J.G.); Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, Scotland, UK (J.E.H., V.V., A.F.W., C.H.); Department of Neurology, Innsbruck Medical University, Innsbruck, Austria (S.K., J.W.); Endocrinology, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD (M.E.M., Q.G., J.O., B.D.M., A.R.S.); Unit of General Epidemiology and Bioinformatics, Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands (I.M.N., H.S.); Department of Epidemiology and Prevention, Wake Forest University School of Medicine, Winston-Salem, NC (M.E.R., Y.L.); Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-Univ Greifswald, Greifswald, Germany (A.T., U.V.); Department of Cardiology (P.H., W.H.G.), and the Department of Genetics, University Medical Center Groningen, Univ of Groningen, Groningen, The Netherlands (P.H.); Hudson Alpha Institute for Biotechnology, Huntsville, AL (L.L.W., D.M.A.); Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, UK (A.R.W., A.M.); Department of Family and Preventive Medicine, University of California San Diego, Preventive Medicine, La Jolla, CA (C.L.W., M.A.A., M.H.C.); Department of Biostatistics, University of Washington, Seattle, WA (A.A., K.R.); Department of Cardiology, Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands (F.W.A.); Julius Center for Health Sciences and Primary Care (F.W.A.), and the Department of Medical Genetics, Biomedical Genetics, University Medical Center, Utrecht, The Netherlands (F.W.A.); Geriatric Rehabilitation Unit, Azienda Sanitaria di Firenze, Florence, Italy (S.B.); University of Texas Health Science Center at Houston, Department of Epidemiology, Human Genetics and Environmental Sciences, Houston, TX (M.B., E.B.); Department of Pharmacology, University of Split, Croatia (M.B., G.G., I.M.); The Center for Human Genetics Research, Vanderbilt University, Nashville, TN (K.B-G., R.G.); Department of Biostatistics, University of North Carolina Chapel Hill, Chapel Hill, NC (D.J.C.); Department of Endocrinology and Epidemiology, Biostatistics and Health Technology Assessment, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands (M.H.); Department of Laboratory Medicine, Konventhospital Barmherzige Brueder Linz, Linz, Austria (B.D., M.H., T.M.); Sticht Center on Aging, Wake Forest School of Med, Winston-Salem, NC (J.D., S.B.K.); Department of Internal Med B- Cardiology, Angiology & Pneumology & Intensive Care Medicine, University Medicine, Greifswald, Germany (M.D., S.B.F., A.K.); Department of Medicine 2, University Medical Center Mainz, Johannes Gutenberg-University Mainz, Germany (C.E-K., K.J.L., P.S.W.); Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, National Institute of Health, Baltimore, MD (L.F.); Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN (A.R.F.); Department of Vascular Surgery, Innsbruck Medical University, Innsbruck, Austria (G.F., B.R.); Department of Epidemiology, Biostatistics & HTA, Radboud Univ Nijmegen Med Center, Nijmegen, The Netherlands (L.A.K., S.H.V.); Department of Public Health, University of Split School of Medicine, Croatia (I.K., O.P.); Cardiovascular Diseases and the Gonda Vascular Center, Mayo Clinic, Rochester, MN (I.J.K.); Medical Genetics Institute, Cedars-Sinai Med Center, Los Angeles, CA (X.L.); Institute for Community Medicine, University Medicine Greifswald, Germany (W.L.); Department of Biostatistics, Wake Forest University School of Medicine, Winston-Salem, NC (K.L.); Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Neuherberg, Germany (C.M.); De-

partment of Epidemiology and Public Health, Peninsula College of Medicine and Dentistry, University of Exeter, UK (D.M.); Perelman School of Medicine at the University of Pennsylvania, Cardiovascular Division, Vascular Medical Section, Philadelphia, PA (E.R.M.); Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, Netherlands (G.N.); Department of Internal Medicine, Bruneck Hospital, Bruneck, Italy (F.O.); Mount Sinai Medical Center, New York, NY (J.W.O.); National Heart, Lung, and Blood Institute, Intramural Research, Bethesda, MD (C.J.O.); Department of Medicine, Columbia University, New York, NY (W.P.); Department of Psychiatry/EMGO Institute, VU University Medical Center, Amsterdam (B.W.P., A.S.); Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands (B.W.P.); Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands (B.W.P.); Institute of Clinical Chemistry and Laboratory Medicine, University Medicine, Greifswald, Germany (A.P.); Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington (B.M.P.); Group Health Research Institute, Group Health Cooperative, Seattle, WA (B.M.P.); Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands (F.R., A.G.U.); University of California Los Angeles, Los Angeles, CA (J.I.R.); Hietzing Hospital, Third Medical Department of Metabolic Diseases and Nephrology, Vienna, Austria (M.S.); Clinical Research Branch, National Institute on Aging, Baltimore, MD (T.T.); Center for Population Health Sciences, University of Edinburgh, Edinburgh, Scotland (S.H.W., L.Z., H.C., I.R., J.F.W.); Center for Thrombosis and Hemostasis, University Medical Center Mainz, Johannes Gutenberg-University Mainz, Germany (P.S.W.); Clinic for General and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany (T.Z., S.B.); Department of Biology, University of Split, Croatia (T.Z.); Andrija Stampar School of Public Health, Medical School, University of Zagreb, Croatia (L.Z., I.R.); Department of Medicine, Stanford University School of Medicine, Stanford, CA (T.L.A., J.P.C.); Department of Internal Med, Wake Forest University School of Medicine, Winston-Salem, NC (D.H.); Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI (S.L.R.K.); Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, PA (A.B.N.); Department of Clinical Genetics, Erasmus Med Center, Rotterdam, The Netherlands (B.A.O.); Geriatric Research and Education Clinical Center, VA Med Center, Baltimore, MD (A.R.S.); Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Neuherberg, Germany (H.E.W.); Carolina Center for Genome Sciences, School of Public Health, University of North Carolina-Chapel Hill, Chapel Hill, NC (K.E.N.).

Acknowledgments

See online-only Data Supplement Material.

Sources of Funding

The Amish Study was supported by grants R01 088119, R01 AR046838, U01 HL72515, and R01 AG18728, the University of Maryland General Clinical Research Center, Grant M01 RR 16500, Mid-Atlantic Nutrition Obesity Research Center Grant P30 DK072488, General Clinical Research Centers Program, National Center for Research Resources (NCRR), and the Baltimore Veterans Administration Geriatric Research and Education Clinical Center (GRECC). Dr. Montasser was supported by AG000219. ARIC is carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and NIH contract HHSN268200625226C and infrastructure support UL1RR025005. CHS research was supported by NHLBI contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, and NHLBI grants HL080295, HL075366, HL087652, HL105756, with additional contribution from NINDS. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIA. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01-RR00425 and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491. The Family Heart Study GWAS was funded by grant HL08770002, and the work was supported by NHLBI contract numbers R01HL08770003I, and R01DK06833603 and R01DK07568101 from NIDDK. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc, for genotyping services (Contract No. N02-HL-6-4278). A portion of this research used the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. "Genetic Epidemiology Network of Arteriopathy (GENOA) study is supported by the NIH, grant number 5R01HL087660." The Gutenberg Heart Study is funded through Rheinland-Pfalz ("Stiftung Rheinland Pfalz für Innovation," contract number AZ 961-386261/733), the research programs "Wissen schafft Zukunft" and "Schwerpunkt Vaskuläre Prävention" of the Johannes Gutenberg-University of Mainz and its contract with Boehringer Ingelheim and PHILIPS Medical Systems, including an unrestricted grant for the Gutenberg Heart Study. This research also was supported by the National Genome Network "NGFNplus" (contract number project A3 01GS0833 and 01GS0831), by the Federal Ministry of Education and Research, Germany. Health ABC was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The GWAS was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences, and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. The InCHIANTI study baseline (1998–2000) was supported as a "targeted project"(ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the NIA (Contracts: 263 MD 9164 and 263 MD 821336); the InCHIANTI Follow-up 1 (2001–2003) was funded by NIA (Contracts: N.1-AG-1-1 and N.1-AG-1-2111; the InCHIANTI Follow-ups 2 and 3 studies (2004–2010) were funded by NIA (Contract: N01-AG-5-0002), supported in part by the NIA Intramural research program. A portion of the support was through a R and D contract with MedStar Health Research Institute. KORA F3 and KORA F4 were partially funded by the "Genomics of Lipid-associated Disorders-GOLD" of the "Austrian Genome Research Programme GEN-AU" and by the Austrian Heart Fund to F. Kronenberg and by the Austrian National Bank (Project-Nr. 13662) to Barbara Kollerits. The MONICA/KORA Augsburg cohort study was financed by the Helmholtz Zentrum München and the German National Genome Research Net NGFN2 and NGFNplus (H.-E. Wichmann 01GS0823). NESDA was supported by the Geestkracht program of ZonMW [grant 10-000-1002]; matching funds from universities and mental health care institutes involved in NESDA (GGZ Buitenamstel-Geestgronden, Rivierduinen, University Medical Center Groningen, GGZ 25 Lentis, GGZ Friesland, GGZ Drenthe). Genotyping was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US NIH, and analysis was supported by grants from GAIN and the NIMH (MH081802). NBS support was obtained from RUNMC. The measurement of ABI was supported by Grant 2003B057 of the Netherlands Heart Foundation. This work was sponsored by the Stichting Nationale Computerfaciliteiten (National Computing Facilities Foundation, NCF) for the use of supercomputer facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research, NWO). The Rotterdam GWA study was funded by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Consortium for Healthy Aging (NCHA) project nr. 050-060-810. Dr Jacqueline Witteman is supported by Netherlands Organization for Scientific Research (NOW) grant (vici, 918-76-619).

Abbas Dehghan is supported by Erasmus University Rotterdam (EUR) fellowship. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. SHIP is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs, as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg-West Pomerania. CROATIA Studies (Croatia-Vis, Croatia-Korcula, and Croatia-Split) were supported by grants from the Medical Research Council UK and Ministry of Science, Education and Sport of the Republic of Croatia (No. 108-1080315-0302), and CROATIA-Vis by the European Union framework program 6 European Special Populations Research Network (EUROSPAN) project (contract LSHG-CT-2006-018947). The genotyping for the ERF study was supported by EUROSPAN (European Special Populations Research Network) and the European Commission FP6 STRP grant (018947; LSHG-CT-2006-01947). The ERF study was further supported by grants from the Netherlands Organisation for Scientific Research, Erasmus MC, the Centre for Medical Systems Biology (CMSB), and the Netherlands Brain Foundation (HersenStichting Nederland). ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. The Bruneck Study was supported by the Pustertaler Verein zur Prävention von Herz- und Hirngefaesserkrankungen, Gesundheitsbezirk Bruneck, and the Assessorat fuer Gesundheit, Province of Bolzano, Italy. The Copenhagen City Heart Study was supported by a Specific Targeted Research Project grant from the European Union, Sixth Framework Programme Priority (FP-2005-LIFESCIHEALTH-6) contract 037631, the Danish Medical Research Council (Copenhagen), the Research Fund at Rigshospitalet, Copenhagen University Hospital (Copenhagen), Chief Physician Johan Boserup and Lise Boserup's Fund (Copenhagen), Ingeborg and Leo Dannin's Grant (Copenhagen), and Henry Hansen's and Wife's Grant (Copenhagen). Genotyping was supported by a grant from the Austrian Heart Fund. National Health and Nutrition Examination Surveys (NHANES) are supported by the Centers for Disease Control and Prevention. The Vanderbilt University Center for Human Genetics Research, Computational Genomics Core, provided computational or analytical support for this work. MESA and the MESA SHARe project are conducted and supported by the NHLBI in collaboration with MESA investigators. Support is provided by grants and contracts N01 HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and RR-024156. "Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278." The PREVEND study is supported by the Dutch Kidney Foundation (Grant E033), EU project grant GENECURE (FP-6 LSHM CT 2006 037697), and NWO VENI (grant number 916.76.170). The CAVASIC study was partially funded by the "Genomics of Lipid-associated Disorders-GOLD" of the "Austrian Genome Research Programme GEN-AU" and by the Austrian Heart Fund and the Austrian National Bank (Project-Nr. 13662). GenePAD was supported by grants RO1 HL-75774, 1K12 HL087746, 1P50HL083800, as well as grant M01 RR 00070 (General Clinical Research Center, Stanford University School of Medicine) and the Stanford Cardiovascular Institute. The LIPAD project was supported in part by a grant from the Upper Austrian Government. Genotyping was supported by a grant from the Austrian Heart Fund.

Disclosures

None.

References

1. Belch JJ, Topol EJ, Agnelli G, Bertrand M, Califf RM, Clement DL, Creager MA, Easton JD, Gavin JR 3rd, Greenland P, Hankey G, Hanrath P, Hirsch AT, Meyer J, Smith SC, Sullivan F, Weber MA. Critical issues in peripheral arterial disease detection and management: a call to action. *Arch Intern Med*. 2003;163:884–892.
2. Criqui MH, Langer RD, Fronek A, Feigelson HS, Klauber MR, McCann TJ, Browner D. Mortality over a period of 10 years in patients with peripheral arterial disease. *N Engl J Med*. 1992;326:381–386.
3. Newman AB, Shemanski L, Manolio TA, Cushman M, Mittelmark M, Polak JF, Powe NR, Siscovick D. Ankle-arm index as a predictor of cardiovascular disease and mortality in the Cardiovascular Health Study. The Cardiovascular Health Study Group. *Arterioscler Thromb Vasc Biol*. 1999;19:538–545.
4. Murabito JM, Evans JC, Larson MG, Nieto K, Levy D, Wilson PW. The ankle-brachial index in the elderly and risk of stroke, coronary disease, and death: the Framingham Study. *Arch Intern Med*. 2003;163:1939–1942.
5. Weatherley BD, Nelson JJ, Heiss G, Chambless LE, Sharrett AR, Nieto FJ, Folsom AR, Rosamond WD. The association of the ankle-brachial index with incident coronary heart disease: the Atherosclerosis Risk In Communities (ARIC) study, 1987–2001. *BMC Cardiovasc Disord*. 2007;7:3.
6. Lamina C, Meisinger C, Heid IM, Lowel H, Rantner B, Koenig W, Kronenberg F; Kora Study Group. Association of ankle-brachial index and plaques in the carotid and femoral arteries with cardiovascular events and total mortality in a population-based study with 13 years of follow-up. *Eur Heart J*. 2006;27:2580–2587.
7. Hirsch AT, Haskal ZJ, Hertzer NR, Bakal CW, Creager MA, Halperin JL, Hiratzka LF, Murphy WR, Olin JW, Puschett JB, Rosenfield KA, Sacks D, Stanley JC, Taylor LM Jr, White CJ, White J, White RA, Antman EM, Smith SC Jr, Adams CD, Anderson JL, Faxon DP, Fuster V, Gibbons RJ, Halperin JL, Hiratzka LFHunt SA, Jacobs AK, Nishimura R, Ornato JP, Page RL, Riegel B. ACC/AHA 2005 Guidelines for the Management of Patients With Peripheral Arterial Disease (Lower Extremity, Renal, Mesenteric, and Abdominal Aortic): A Collaborative Report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease). *J Am Coll Cardiol*. 2006;47:e1–e192.
8. Fowkes FG, Murray GD, Butcher I, Heald CL, Lee RJ, Chambless LE, Folsom AR, Hirsch AT, Dramaix M, deBacker G, Wautrecht JC, Kornitzer M, Newman AB, Cushman M, Sutton-Tyrrell K, Fowkes FG, Lee AJ, Price JF, d'Agostino RB, Murabito JM, Norman PE, Jamrozik K, Curb JD, Masaki KH, Rodríguez BL, Dekker JM, Bouter LM, Heine RJ, Nijpels G, Stehouwer CD, Ferrucci L, McDermott MM, Stoffers HE, Hooi JD, Knottnerus JA, Ogren M, Hedblad B, Witteman JC, Breteler MM, Hunink MG, Hofman A, Criqui MH, Langer RD, Fronek A, Hiatt WR, Hamman R, Resnick HE, Guralnik J, McDermott MM. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA*. 2008;300:197–208.
9. Valentine RJ, Guerra R, Stephan P, Scoggins E, Clagett GP, Cohen J. Family history is a major determinant of subclinical peripheral arterial disease in young adults. *J Vasc Surg*. 2004;39:351–356.
10. Kullo IJ, Turner ST, Kardia SL, Mosley TH, Jr., Boerwinkle E, de Andrade M. A genome-wide linkage scan for ankle-brachial index in African American and non-Hispanic white subjects participating in the GENOA study. *Atherosclerosis*. 2006;187:433–438.
11. Murabito JM, Guo CY, Fox CS, D'Agostino RB. Heritability of the ankle-brachial index: the Framingham Offspring study. *Am J Epidemiol*. 2006;164:963–968.
12. Carmelli D, Fabsitz RR, Swan GE, Reed T, Miller B, Wolf PA. Contribution of genetic and environmental influences to ankle-brachial blood pressure index in the NHLBI Twin Study. National Heart, Lung, and Blood Institute. *Am J Epidemiol*. 2000;151:452–458.
13. Knowles JW, Assimes TL, Li J, Quertermous T, Cooke JP. Genetic susceptibility to peripheral arterial disease: a dark corner in vascular biology. *Arterioscler Thromb Vasc Biol*. 2007;27:2068–2078.
14. Kardia SL, Greene MT, Boerwinkle E, Turner ST, Kullo IJ. Investigating the complex genetic architecture of ankle-brachial index, a measure of

peripheral arterial disease, in non-Hispanic whites. *BMC Med Genomics*. 2008;1:16.
15. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H; WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357:443–453.
16. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007;316:1488–1491.
17. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, König IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ; Cardiogenics Consortium. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation*. 2008;117:1675–1684.
18. Shen GQ, Li L, Rao S, Abdullah KG, Ban JM, Lee BS, Park JE, Wang QK. Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2008; 28:360–365.
19. Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, Meitinger T, Wichmann E, Mitchell BD, Furie K, Slowik A, Rich SS, Syme PD, MacLeod MJ, Meschia JF, Rosand J, Kittner SJ, Markus HS, Müller-Myhsok B, Dichgans M; International Stroke Genetics Consortium. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. *Ann Neurol*. 2009;65:531–539.
20. Helgadottir A, Thorleifsson G, Magnusson KP, Gretarsdottir S, Steinthorsdottir V, Manolescu A, Jones GT, Rinkel GJ, Blankensteijn JD, Ronkainen A, Jääskeläinen JE, Kyo Y, Lenk GM, Sakalihasan N, Kostulas K, Gottsäter A, Flex A, Stefansson H, Hansen T, Andersen G, Weinsheimer S, Borch-Johnsen K, Jorgensen T, Shah SH, Quyyumi AA, Granger CB, Reilly MP, Austin H, Levey AI, Vaccarino V, Palsdottir E, Walters GB, Jonsdottir T, Snorradottir S, Magnusdottir D, Gudmundsson G, Ferrell RE, Sveinbjornsdottir S, Hernesniemi J, Niemelä M, Limet R, Andersen K, Sigurdsson G, Benediktsson R, Verhoeven EL, Teijink JA, Grobbee DE, Rader DJ, Collier DA, Pedersen O, Pola R, Hillert J, Lindblad B, Valdimarsson EM, Magnadottir HB, Wijmenga C, Tromp G, Baas AF, Ruigrok YM, van Rij AM, Kuivaniemi H, Powell JT, Matthiasson SE, Gulcher JR, Thorgeirsson G, Kong A, Thorsteinsdottir U, Stefansson K. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet*. 2008;40:217–224.
21. Thompson AR, Golledge J, Cooper JA, Hafez H, Norman PE, Humphries SE. Sequence variant on 9p21 is associated with the presence of abdominal aortic aneurysm disease but does not have an impact on aneurysmal expansion. *Eur J Hum Genet*. 2009;17:391–394.
22. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JC, Boerwinkle E; CHARGE Consortium. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2:73–80.
23. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007;449:851–861.
24. Servin B, Stephens M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet*. 2007;3:e114.
25. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39:906–913.
26. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–2191.
27. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009;106:9362–9367.
28. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol*. 2008;32:381–385.
29. Johnson AD, O'Donnell CJ. An open access database of genome-wide association results. *BMC Med Genet*. 2009;10:6.
30. Ziegler A, Van Steen K, Wellek S. Investigating Hardy-Weinberg equilibrium in case-control or cohort studies or meta-analysis. *Breast Cancer Res Treat*. 2011;128:197–201.
31. Gao X. Multiple testing corrections for imputed SNPs. *Genet Epidemiol*. 2011;35:154–158.
32. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Coronary Artery Disease (C4D) Genetics Consortium. *Nat Genet.* 2011;43:339–344.
33. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I; Cardiogenics, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C,

Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J; CARDIoGRAM Consortium, Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43:333–338.

34. Wang F, Xu CQ, He Q, Cai JP, Li XC, Wang D, Xiong X, Liao YH, Zeng QT, Yang YZ, Cheng X, Li C, Yang R, Wang CC, Wu G, Lu QL, Bai Y, Huang YF, Yin D, Yang Q, Wang XJ, Dai DP, Zhang RF, Wan J, Ren JH, Li SS, Zhao YY, Fu FF, Huang Y, Li QX, Shi SW, Lin N, Pan ZW, Li Y, Yu B, Wu YX, Ke YH, Lei J, Wang N, Luo CY, Ji LY, Gao LJ, Li L, Liu H, Huang EW, Cui J, Jia N, Ren X, Li H, Ke T, Zhang XQ, Liu JY, Liu MG, Xia H, Yang B, Shi LS, Xia YL, Tu X, Wang QK. Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nat Genet*. 2011;43:345–349.
35. Cluett C, McDermott MM, Guralnik J, Ferrucci L, Bandinelli S, Miljkovic I, Zmuda JM, Li R, Tranah G, Harris T, Rice N, Henley W, Frayling TM, Murray A, Melzer D. The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people. *Circ Cardiovasc Genet*. 2009;2:347–353.
36. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007;316:1491–1493.
37. Cunnington MS, Mayosi BM, Hall DH, Avery PJ, Farrall M, Vickers MA, Watkins H, Keavney B. Novel genetic variants linked to coronary artery disease by genome-wide association are not associated with carotid artery intima-media thickness or intermediate risk phenotypes. *Atherosclerosis*. 2009;203:41–44.
38. Samani NJ, Raitakari OT, Sipila K, Tobin MD, Schunkert H, Juonala M, Braund PS, Erdmann J, Viikari J, Moilanen L, Taittonen L, Jula A, Jokinen E, Laitinen T, Hutri-Kähönen N, Nieminen MS, Kesäniemi YA, Hall AS, Hulkkonen J, Kähönen M, Lehtimäki T. Coronary artery disease-associated locus on chromosome 9p21 and early markers of atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2008;28:1679–1683.
39. Ye S, Willeit J, Kronenberg F, Xu Q, Kiechl S. Association of genetic variation on chromosome 9p21 with susceptibility and progression of atherosclerosis: a population-based, prospective study. *J Am Coll Cardiol*. 2008;52:378–384.
40. Lin HF, Tsai PC, Lin RT, Khor GT, Sheu SH, Juo SH. Sex differential genetic effect of chromosome 9p21 on subclinical atherosclerosis. *PLoS One*. 2010;5:e15124.
41. Bhatt DL, Steg PG, Ohman EM, Hirsch AT, Ikeda Y, Mas JL, Goto S, Liau CS, Richard AJ, Röther J, Wilson PW; REACH Registry Investigators. International prevalence, recognition, and treatment of cardiovascular risk factors in outpatients with atherothrombosis. *JAMA*. 2006;295:180–189.
42. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Boström KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jørgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjögren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ; Wellcome Trust Case Control Consortium, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. 2008;40:638–645.
43. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Boström K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Råstam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjögren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. 2007;316:1331–1336.
44. Caporaso N, Gu F, Chatterjee N, Sheng-Chih J, Yu K, Yeager M, Chen C, Jacobs K, Wheeler W, Landi MT, Ziegler RG, Hunter DJ, Chanock S, Hankinson S, Kraft P, Bergen AW. Genome-wide and candidate gene association study of cigarette smoking behaviors. *PLoS One*. 2009;4:e4653.
45. Musunuru K, Post WS, Herzog W, Shen H, O'Connell JR, McArdle PF, Ryan KA, Gibson Q, Cheng YC, Clearfield E, Johnson AD, Tofler G, Yang Q, O'Donnell CJ, Becker DM, Yanek LR, Becker LC, Faraday N, Bielak LF, Peyser PA, Shuldiner AR, Mitchell BD. Association of single nucleotide polymorphisms on chromosome 9p21.3 with platelet reactivity: a potential mechanism for increased vascular disease. *Circ Cardiovasc Genet*. 2010;3:445–453.
46. Cunnington MS, Keavney B. Genetic mechanisms mediating atherosclerosis susceptibility at the chromosome 9p21 locus. *Curr Atheroscler Rep*. 2011;13:193–201.
47. Holdt LM, Sass K, Gabel G, Bergert H, Thiery J, Teupser D. Expression of Chr9p21 genes CDKN2B (p15(INK4b)), CDKN2A (p16(INK4a), p14(ARF)) and MTAP in human atherosclerotic plaque. *Atherosclerosis*. 2011;214:264–270.
48. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature*. 2010;464:409–412.
49. Pasmant E, Sabbagh A, Vidaud M, Bieche I. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB J*. 2011;25:444–448.
50. Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG, Frazer KA. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature*. 2011;470:264–268.
51. Gretarsdottir S, Baas AF, Thorleifsson G, Holm H, den Heijer M, de Vries JP, Kranendonk SE, Zeebregts CJ, van Sterkenburg SM, Geelkerken RH, van Rij AM, Williams MJ, Boll AP, Kostic JP, Jonasdottir A, Jonasdottir A, Walters GB, Masson G, Sulem P, Saemundsdottir J, Mouy M, Magnusson KP, Tromp G, Elmore JR, Sakalihasan N, Limet R, Defraigne JO, Ferrell RE, Ronkainen A, Ruigrok YM, Wijmenga C, Grobbee DE, Shah SH, Granger CB, Quyyumi AA, Vaccarino V, Patel RS, Zafari AM, Levey AI, Austin H, Girelli D, Pignatti PF, Olivieri O, Martinelli N, Malerba G, Trabetti E, Becker LC, Becker DM, Reilly MP, Rader DJ, Mueller T, Dieplinger B, Haltmayer M, Urbonavicius S, Lindblad B, Gottsäter A, Gaetani E, Pola R, Wells P, Rodger M, Forgie M, Langlois N, Corral J, Vicente V, Fontcuberta J, España F, Grarup N, Jørgensen T, Witte DR, Hansen T, Pedersen O, Aben KK, de Graaf J, Holewijn S, Folkersen L, Franco-Cereceda A, Eriksson P, Collier DA, Stefansson H, Steinthorsdottir V, Rafnar T, Valdimarsson EM, Magnadottir HB, Sveinbjornsdottir S, Olafsson I, Magnusson MK, Palmason R, Haraldsdottir V, Andersen K, Onundarson PT, Thorgeirsson G, Kiemeney LA, Powell JT, Carey DJ, Kuivaniemi H, Lindholt JS, Jones GT, Kong A, Blankensteijn JD, Matthiasson SE, Thorsteinsdottir U, Stefansson K. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. *Nat Genet*. 2010;42:692–697.
52. Harrison SC, Holmes MV, Agu O, Humphries SE. Genome wide association studies of abdominal aortic aneurysms-biological insights and potential translation applications. *Atherosclerosis*. 2011;217:47–56.
53. Renner W, Schallmoser K, Gallippi P, Krauss C, Toplak H, Wascher TC, Pilger E. C242T polymorphism of the p22 phox gene is not associated with peripheral arterial occlusive disease. *Atherosclerosis*. 2000;152:175–179.
54. Linsel-Nitschke P, Gotz A, Erdmann J, Braenne I, Braund P, Hengstenberg C, Stark K, Fischer M, Schreiber S, El Mokhtari NE, Schaefer A, Schrezenmeir J, Rubin D, Hinney A, Reinehr T, Roth C, Ortlepp J, Hanrath P, Hall AS, Mangino M, Lieb W, Lamina C, Heid IM, Doering A, Gieger C, Peters A, Meitinger T, Wichmann HE, König IR, Ziegler A, Kronenberg F, Samani NJ, Schunkert H; Wellcome Trust Case Control Consortium (WTCCC); Cardiogenics Consortium. Lifelong reduction of

LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease–a Mendelian Randomisation study. *PLoS One*. 2008;3:e2986.
55. Allison MA, Criqui MH, McClelland RL, Scott JM, McDermott MM, Liu K, Folsom AR, Bertoni AG, Sharrett AR, Homma S, Kori S. The effect of novel cardiovascular risk factors on the ethnic-specific odds for peripheral arterial disease in the Multi-Ethnic Study of Atherosclerosis (MESA). *J Am Coll Cardiol*. 2006;48:1190–1197.
56. Conen D, Everett BM, Kurth T, Creager MA, Buring JE, Ridker PM, Pradhan AD. Smoking, smoking cessation [corrected], and risk for symptomatic peripheral artery disease in women: a cohort study. *Ann Intern Med*. 2011;154:719–726.
57. Viiri LE, Viiri KM, Ilveskoski E, Huhtala H, Maki M, Tienari PJ, Perola M, Lehtimäki T, Karhunen PJ. Interactions of functional apolipoprotein E gene promoter polymorphisms with smoking on aortic atherosclerosis. *Circ Cardiovasc Genet*. 2008;1:107–116.

CLINICAL PERSPECTIVE

Little is known about the genetic susceptibility to peripheral arterial disease (PAD). We conducted a meta-analysis of genome-wide association study findings for the ankle-brachial index (ABI), a noninvasive measure of PAD, within an international consortium of 21 population-based cohort studies that included over 40 000 participants of European descent, and conducted replication analyses in over 16 000 individuals from population-based cohorts and clinically-based studies of PAD. We identified and replicated 1 genome-wide significant association between a genetic variant in the chromosome 9p21 region and a lower ABI. Common genetic variants in the 9p21 locus are associated strongly with coronary artery disease and confer risk for other atherosclerotic diseases. Therefore, the primary effect of the 9p21 region may be on the atherosclerotic process itself, and there are likely many other factors, both genetic and environmental, that determine whether it manifests as coronary disease, PAD, or another clinical atherosclerotic phenotype. The primary biological mechanism underlying the association with ABI is unknown but appears independent of 2 major PAD risk factors, diabetes and smoking, as the ABI single nucleotide polymorphisms (SNP) in the 9p21 region we identified is not in linkage disequilibrium with the SNPs in the region associated with diabetes or smoking-related behaviors. PAD represents a diffuse form of atherosclerosis associated with increased risk for death and incident CVD events. Identification of genetic variants associated with ABI may provide an opportunity to unravel the biological basis of PAD.

Supplemental Material

Association between chromosome 9p21 variants and the Ankle-brachial Index identified by a meta-analysis of 21 genome-wide association studies

Joanne M. Murabito, MD, ScM*[1,2], Charles C. White, MPH*[3], Maryam Kavousi, MD, MSc*[4,5], Yan V. Sun, PhD*[6], Mary F. Feitosa, PhD*[7], Vijay Nambi, MD*[8], Claudia Lamina, PhD*[9], Arne Schillert, PhD*[10], Stefan Coassin, PhD[9], Joshua C. Bis, PhD[11], Linda Broer, MSc[4,5], Dana C. Crawford, PhD[12], Nora Franceschini, MD, MPH[13], Ruth Frikke-Schmidt, MD, PhD[14], Margot Haun, MSc[9], Suzanne Holewijn, PhD[15], Jennifer E. Huffman, MSc[16], Shih-Jen Hwang, PhD[1], Stefan Kiechl, MD[17], Barbara Kollerits, PhD, MPH[9], May E. Montasser, PhD[18], Ilja M. Nolte, PhD[19], Megan E. Rudock, PhD[20], Andrea Senft, MSc[10], Alexander Teumer, PhD[21], Pim van der Harst, MD, PhD[22,23], Veronique Vitart, PhD[16], Lindsay L. Waite, MS[24], Andrew R. Wood, MRes[25], Christina L. Wassel, PhD[26], Devin M. Absher, PhD[24], Matthew A. Allison, MD, MPH[26], Najaf Amin, PhD[4], Alice Arnold, PhD[27], Folkert W. Asselbergs MD, PhD[28,29,30], Yurii Aulchenko, PhD[4], Stefania Bandinelli, MD[31], Maja Barbalic, PhD[32], Mladen Boban, MD, PhD[33], Kristin Brown-Gentry, MS[34], David J. Couper, PhD[35], Michael H. Criqui, MD, MPH[26], Abbas Dehghan, MD, PhD[4,5], Martin den Heijer, MD, PhD[36], Benjamin Dieplinger, MD[37], Jingzhong Ding, PhD[38], Marcus Dörr, MD[39], Christine Espinola-Klein, MD[40], Stephan B. Felix, MD[39], Luigi Ferrucci, MD, PhD[41], Aaron R. Folsom, MD[42], Gustav Fraedrich, MD[43], Quince Gibson, MBA[18], Robert Goodloe, MS[34], Grgo Gunjaca, MD[33], Meinhard Haltmayer, MD[37], Gerardo Heiss, MD, PhD[13], Albert Hofman, MD, PhD[4,5], Arne Kieback, MD[39], Lambertus A. Kiemeney, PhD[44], Ivana Kolcic, MD, PhD[45], Iftikhar J. Kullo, MD[46], Stephen B. Kritchevsky, PhD[38], Karl J. Lackner, MD[40], Xiaohui Li, MD, MSc[47], Wolfgang Lieb, MD, MSc[48], Kurt Lohman, Mstat[49], Christa Meisinger, MD, MPH[50], David Melzer, MD, PhD[51], Emile R Mohler III, MD[52], Ivana Mudnic, MD[33], Thomas Mueller, MD[37], Gerjan Navis, MD, PhD[53], Friedrich Oberhollenzer, MD[54], Jeffrey W. Olin, MD[55], Jeff O'Connell, PhD[18], Christopher J. O'Donnell, MD, MPH[1,56], Walter Palmas, MD, MS[57], Brenda W. Penninx, PhD[58,59,60], Astrid Petersmann, MD, PhD[61], Ozren Polasek, MD, PhD[45], Bruce M. Psaty, MD, PhD[62,63], Barbara Rantner, MD, PhD[9,43], Ken Rice, PhD[27], Fernando Rivadeneira, MD, PhD[4,5,64], Jerome I. Rotter, MD[65], Adrie Seldenrijk, PhD[58], Marietta Stadler, MD[66], Monika Summerer, PhD[9], Toshiko Tanaka, PhD[67], Anne Tybjaerg-Hansen, MD, DMSc[14], Andre G. Uitterlinden, PhD[4,5,64], Wiek H. van Gilst, PhD[22], Sita H. Vermeulen, PhD[44], Sarah H. Wild, MB, BChir, PhD[68], Philipp S. Wild, MD[40,69], Johann Willeit, MD[17], Tanja Zeller, PhD[70], Tatijana Zemunik, MD, PhD[71], Lina Zgaga, MD, PhD[68,72], Themistocles L. Assimes, MD, PhD[73], Stefan Blankenberg, MD[70], Eric Boerwinkle, PhD[32], Harry Campbell, MD[68], John P. Cooke, MD, PhD[73], Jacqueline de Graaf, MD, PhD[15], David Herrington, MD, MHS[74], Sharon L. R. Kardia, PhD[75], Braxton D. Mitchell, PhD[18], Anna Murray, PhD[25], Thomas Münzel, MD[40], Anne Newman, MD, MPH[76], Ben A. Oostra, PhD[77], Igor Rudan, MD, PhD, MPH[68,72], Alan R. Shuldiner, MD[18,78], Harold Snieder PhD[19],Cornelia M. van Duijn, PhD[4,5], Uwe Völker, PhD[21], Alan F. Wright, PhD[16], H.-Erich Wichmann, MD, PhD[79], James F. Wilson, DPhil[68], Jacqueline C.M. Witteman, PhD[4,5], Yongmei Liu, MD, PhD*[20], Caroline Hayward, PhD*[16], Ingrid B. Borecki, PhD*[7], Andreas Ziegler, PhD*[10], Kari E. North, PhD*[13,80], L. Adrienne Cupples, PhD*[1,3], Florian Kronenberg, MD*[9].

1. Study Specific Phenotype and Genotype Information

1.1 Discovery cohorts: non-isolated populations

1.1.1 Atherosclerosis Risk in Communities Study (ARIC)

The ARIC study is a predominantly bi-racial prospective study of the etiology and natural history of subclinical and clinical atherosclerosis among 15,792 middle-aged men and women, aged 45 to 64 years old at recruitment in 1987 to 1989 [(1)]. Study participants were selected as a probability sample from four US communities: Forsyth Co, NC; Jackson, MS; seven northwestern suburbs of Minneapolis, MN; and Washington Co, MD. African-Americans were over-sampled in Forsyth CO and were exclusively sampled in Jackson. Participants were examined at baseline with collection of medical, social, and demographic data and at three triennial follow-up exams, the last of which occurred in 1999. Annual follow-up continues to ascertain vital status as well as CHD and stroke, including hospitalizations and deaths. GWAS data are available from 8,861 self-identified white ARIC participants.

1.1.2 Cardiovascular Health Study (CHS)

The CHS is a population-based cohort study of risk factors for CHD and stroke in adults age 65 years and older conducted across four field centers in the United States [(2)]. The original predominantly Caucasian cohort of 5201 persons was recruited in 1989-1990 from a random sample of people on Medicare eligibility lists and an additional 687 African-Americans were enrolled subsequently for a total sample of 5,888.

DNA was extracted from blood samples drawn on all participants at their baseline examination in 1989 to 1990. Genotyping was performed from 2007 to 2008 at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars- Sinai on the 3,980 CHS participants who had consented to genetic testing and were free of CVD at baseline. A total of 1,908 persons were excluded for lack of available DNA or prevalent coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke, or transient ischemic attack. Because the other cohorts were predominantly white, the African-American participants were excluded from this analysis to limit the potential for false positive associations due to population stratification, leaving 3397 Caucasians successfully genotyped, 3291 had measures of ABI.

1.1.3 Family Heart Study (FamHS)

The Family Heart Study (https://dsgweb.wustl.edu/PROJECTS/MP1.html) was begun in 1992 with the ascertainment of 1,200 families, half randomly sampled, and half selected because of an excess of coronary heart disease (CHD) or risk factor abnormalities as compared with age- and sex-specific population rates [(3)]. The families, with approximately 6,000 individuals, were sampled on the basis of information on probands from four population-based parent studies: the Framingham Heart Study, the Utah Family Tree Study, and two ARIC centers (Minneapolis, and Forsyth County, NC). A broad range of phenotypes were assessed at a clinic examination (exam 1) in broad domains of CHD, atherosclerosis, cardiac and vascular function, inflammation and hemostasis, lipids and lipoproteins, blood pressure, diabetes and insulin resistance, pulmonary function, and anthropometry. Approximately 8 years later, study participants belonging to the largest pedigrees were invited for a second clinical exam. A total of 2,767 Caucasian subjects in 510 extended families were examined. This sample is the focus of the genome-wide association study (GWAS).

A two-stage design was adopted for the GWAS. In the first stage, 1016 subjects were chosen, equally distributed between the upper and lower quartile of age- and sex-adjusted values for coronary artery calcification, assessed by CT scan. These subjects were chosen to be largely unrelated; 34% of the subjects were from unique families, while 200 other subjects had 1 or more siblings selected into the sample, yielding a sample of 465 unrelated subjects. The remaining family members (n=1,753) will be genotyped in the second stage for replication of the top hits from the first stage.

1.1.4 Framingham Heart Study (FHS)

The FHS is a National Heart Lung and Blood Institute contract-funded observational cohort study initiated in 1948 to examine the determinants of cardiovascular disease and its risk factors (http://www.framinghamheartstudy.org/). The Original Cohort comprised 5,209 men and women, aged 28-62 years at enrollment who have undergone routine biennial examinations [4,5]. In 1971, 5,124 Offspring of the Original Cohort participants and Offspring spouses, aged 5 to 70 years, were enrolled into the Framingham Offspring Study and have been examined approximately every 4 years [6,7]. In the 1990s, DNA was obtained for genetic studies from surviving Original cohort and Offspring participants. Routine examinations for all FHS cohorts included a standardized physician administered medical history interview and physical examination, direct measurement of cardiovascular risk factors, laboratory assessment, and various questionnaires and noninvasive cardiovascular tests specific to the given examination cycle. Cardiovascular events were adjudicated by a panel of senior investigators using previously reported criteria. FHS examinations were approved by the Institutional Review Board of the Boston University Medical Center and all participants provided written informed consent.

1.1.5 Genetic Epidemiology Network of Arteriopathy Study (GENOA)

The Family Blood Pressure Program (FBPP), established by the National Heart Lung and Blood Institute in 1996, joined existing research networks that were investigating hypertension and cardiovascular diseases (http://public.nhlbi.nih.gov/GeneticsGenomics/home/fbpp.aspx). One of the four FBPP networks is the Genetic Epidemiology Network of Arteriopathy (GENOA), which recruited hypertensive, Caucasian sibships for linkage and association studies to investigate genetic contributions to hypertension and hypertension-related target organ damage. Sibships containing at least two individuals with clinically-diagnosed essential hypertension before age 60 years were recruited from Rochester, Minnesota. After identifying each hypertensive sibship, all members of the sibship were invited to participate regardless of their hypertension status. Informed consent was obtained from all subjects and approval was granted by participating institutional review boards. Participants were diagnosed with hypertension if they had either 1) a previous clinical diagnosis of hypertension by a physician with current anti-hypertensive treatment, or 2) an average systolic blood pressure (SBP) ≥ 140 or diastolic blood pressure (DBP) ≥ 90 on the second and third clinic visit as stipulated by the Joint National Committee-7 guidelines [8]. Exclusion criteria were secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy.

1.1.6 Gutenberg Heart Study (GHS)

The Gutenberg Health Study (GHS) is designed as a population-based, prospective, observational single-center cohort study in the Rhein-Main-Region in western Mid-Germany to evaluate and improve cardiovascular risk stratification. It also investigates cancer, eye diseases, diseases of the immune system

and the metabolism. The sample was drawn randomly from the governmental local registry offices in the city of Mainz and the district of Mainz-Bingen. The sample was stratified 1:1 for gender and for urban and rural residence with equal strata for decades of age. Individuals between 35 and 74 years of age were enrolled, and written informed consent was obtained from all participants. Exclusion criteria were insufficient knowledge of the German language, and physical or psychological inability to participate in the examinations at the study center. All participants underwent a five hour examination including a biobanking in the study center. Follow-up examinations are carried out after 2.5 and 5 years. The study protocol was approved by the local ethics committee and the local data safety commissioner and the sampling design by the federal data safety commissioners. All subjects gave written informed consent.

1.1.7 Health, Aging, and Body Composition(Health ABC) Study

The Health Aging and Body Composition (Health ABC) Study is a NIA-sponsored ongoing cohort study of the factors that contribute to incident disability and the decline in function of healthier older persons, with a particular emphasis on changes in body composition in old age. Health ABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. The key components of Health ABC include a baseline exam, annual follow-up clinical exams, and phone contacts every 6 months to identify major health events and document functional status between clinic visits. GWAS data are available from 1663 white participants.

1.1.8 Invecchiare in Chianti (InCHIANTI) Study

The InCHIANTI study is a population based longitudinal study designed to investigate the causes of decline in mobility in older subjects [(9)]. The study sample comprises 1,453 individuals of white European descent and is representative of the older population aged 65 years and older supplemented with at least 30 men and 30 women for each 10 year age group from 20 to 70 years from two small towns from the Chianti region in Tuscany, Italy. Interviews were conducted at the participants' homes by three experienced interviewers. Within three weeks, participants attended a study clinic for blood drawing tests, having fasted for at least 8 and just concluded the 24-hour urine collection. On the same day, the participants received a series of medical examinations, including an assessment of the ankle-brachial index. On the second appointment a clinical evaluation and a comprehensive motor and cognitive performances session were performed by an experienced geriatrician and a trained physical therapist, respectively. INRCA Ethical Committee approved the entire study protocol.

1.1.9 and 1.1.10 Cooperative Health Research in the Region of Augsburg, Kooperative Gesundheitsforschung in der region Augsburg (KORA F3, KORA F4)

The KORA surveys S3 and S4 (Kooperative Gesundheitsforschung in der Region Augsburg [(10)]) are population-based samples drawn from the general population of the South-German city of Augsburg and surrounding counties stratified by age (25 to 74 years) and sex. The follow-ups of these surveys were conducted in 2004/05 (KORA F3) and 2006-2008 (KORA F4). All participants had a German passport, were of European origin and underwent a standardized face-to-face interview by certified medical staff and a standardized medical examination. Genome-wide data are available for a subset of 1,644 subjects randomly chosen from

KORA F3 and for 1814 subjects from KORA F4. ABI-measures were additionally available for n=1,581 (KORA F3) and n=1,407 (KORA F4).

1.1.11. Netherlands Study of Depression and Anxiety (NESDA)

The Netherlands Study of Depression and Anxiety (NESDA) [11], is a multi-centre study designed to examine the long-term course and consequences of depressive and anxiety disorders (http://www.nesda.nl). NESDA included both individuals with depressive and/or anxiety disorders and controls without psychiatric conditions. Inclusion criteria were age 18-65 years and self-reported western European ancestry, exclusion criteria were not being fluent in Dutch and having a primary diagnosis of another psychiatric condition (psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe substance use disorder).

For all participants DNA was isolated from the baseline blood sample. Through funding from the fNIH GAIN program (www.fnih.gov/gain), whole genome scan analysis was conducted for 1,859 NESDA (1,702 depressed cases and 157 controls) participants. A hundred subjects were excluded because of various quality control issues [12].

1.1.12 Nijmegen Biomedical Study (NBS)

The Nijmegen Biomedical Study is a population-based cross-sectional study conducted by the Radboud University Nijmegen Medical Centre [13]. Nijmegen is a town in the eastern part of The Netherlands with 156,000 inhabitants, approximately 87% of Caucasian descent. Age and sex stratified randomly selected adult (age 18 years and older) inhabitants of Nijmegen (N=22,452) received an invitation to fill out a postal questionnaire on lifestyle and medical history. A total of 6,434 participants donated blood for DNA-isolation. GWA-data were available for 1,832 participants [14]. We invited participants aged 50–70 yrs to come to the hospital and performed ABI at rest and after exercise. In total 1517 participants from the Nijmegen Biomedical Study were included [15]. For 544 participants both ABI and GWA-data were available. Approval to conduct the study was obtained from the Institutional Review Board.

1.1.13 and 1.1.14 The Rotterdam Study I and II (RS-I and RS-II)

The Rotterdam Study is a prospective population-based cohort study to investigate the determinants of chronic diseases among participants aged 55 years and older [16]. Briefly, residents of Ommoord, a district of Rotterdam, in the Netherlands, 55 years of age or older, were asked to participate, of whom 7,983 participated (RS-I). The baseline examination was conducted in 1990 -1993 and consisted of a home interview and research center visit for blood samples. In 1999, inhabitants who turned 55 years of age or moved into the study district since the start of the study were invited of whom 3,011 participated (RS-II). In total, at the time of the current Analyses, ABI data were available for 5,169 participants with GWA study data for RS-I and for 1,642 participants from RS-II. The Medical Ethics Committee of Erasmus MC approved the study, and all participants gave informed consent.

1.1.15 Study of Health in Pomerania (SHIP)

The Study of Health in Pomerania (SHIP) is a population-based survey in West Pomerania, the north-east area of Germany [17,18]. A sample from the population aged 20 to 79 years was drawn from population registries. First, the three cities of the region (with 17,076 to 65,977 inhabitants) and the 12 towns (with 1,516 to 3,044 inhabitants) were selected, and then 17 out of 97 smaller towns (with less than 1,500 inhabitants), were drawn at random. Second, from each of the selected communities, subjects were drawn

at random, proportional to the population size of each community and stratified by age and gender. Only individuals with German citizenship and main residency in the study area were included. Finally, 7,008 subjects were sampled, with 292 persons of each gender in each of the twelve five-year age strata. In order to minimize drop-outs by migration or death, subjects were selected in two waves. The net sample (without migrated or deceased persons) comprised 6,267 eligible subjects. Selected persons received a maximum of three written invitations. In case of non-response, letters were followed by a phone call or by home visits if contact by phone was not possible. The SHIP population finally comprised 4,308 participants (corresponding to a final response of 68.7%). Baseline examinations were performed between 1997 and 2001 (SHIP-0). Follow-up examinations were conducted between 2002 and 2006 (SHIP-1) as well as between 2008 and 2012 (SHIP-2). Data for the present analyses were taken from a sample of SHIP-2.

1.2 Discovery cohorts: isolated populations

1.2.1 The Old Order Amish Study (Amish)

The Old Order Amish individuals included in this study were participants of several ongoing studies of cardiovascular health carried out at the University of Maryland [19,20]. Participants were relatively healthy volunteers from the Old Order Amish community of Lancaster County, PA. and their family members. Examinations were conducted at the Amish Research Clinic in Strasburg, PA. All protocols were approved by the Institutional Review Board at the University of Maryland and informed consent was obtained, including permission to use their DNA for genetic studies.

1.2.2, 1.2.3 and 1.2.4 CROATIA-Vis, CROATIA-Korcula and CROATIA-Split

The CROATIA studies recruited adult individuals within each community irrespective of any specific phenotype. Fasting blood samples were collected, biochemical and physiological measurements taken and questionnaire data for medical history as well as lifestyle and environmental exposures were collected following similar protocols. The CROATIA study conformed to the ethical guidelines of the 1975 Declaration of Helsinki and were approved by appropriate ethics boards with all participants signing informed consent prior to participation.

The CROATIA-Vis study includes 1008 Croatians, aged 18–93 years, who were recruited from the villages of Vis and Komiza on the Dalmatian island of Vis during 2003 and 2004 within a larger genetic epidemiology program. The CROATIA-Korcula study includes 969 Croatians between the ages of 18 and 98. The field work was performed in 2007 and 2008 in the eastern part of the island, targeting healthy volunteers from the town of Korčula and the villages of Lumbarda, Žrnovo and Račišće. The CROATIA-Split study includes 499 Croatians between the ages of 18 and 85 recruited in 2009 from the city of Split.

1.2.5 Erasmus Rucphen Family Study (ERF)

The Erasmus Rucphen Family study (ERF) is a cross-sectional family-based study. The study population essentially consists of one extended family of descendents from 20 related couples who lived between 1850 and 1900 and had at least 6 children baptized in the community church. The detailed information about ERF can be found elsewhere [21]. The Medical Ethical Committee of the Erasmus Medical Center, Rotterdam approved the study and informed consent was obtained from all participants.

1.2.6 Orkney Complex Disease Study (ORCADES)

The Orkney Complex Disease Study (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. Data for participants aged 18-91 years, from a subgroup of ten islands, were used for this analysis. Fasting blood samples were collected and over 300 health-related phenotypes and environmental exposures were measured in each individual. All participants gave informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen.

1.3 Replication studies

1.3.1 Bruneck Study

The Bruneck Study is a prospective, population-based survey on the epidemiology, pathophysiology and prevention of cardiovascular and cerebrovascular disease [22-26]. In 1990, 1000 individuals aged 40-79 years were randomly chosen from the inhabitants of the town of Bruneck (South Tyrol, Italy) on the basis of an age- and sex-stratified strategy (125 persons per sex and decade). The population of Bruneck is exclusively Caucasian and of heterogeneous geographic origin with sizeable segments of Austro-German or Italian background. Population mobility within the survey area was low at 0.2% per year. Follow-up examinations were performed in 1995, 2000 and 2005 with participation rates exceeding 90%. Analyses for the current investigation focused on the 1995 evaluation. The study protocol was approved by the pertinent ethics committee, and all participants gave their written informed consent.

1.3.2 Copenhagen City Heart Study (CCHS)

CCHS is a prospective cardiovascular study of the Danish general population initiated in 1976-1978 with follow-up examinations in 1981-1983, 1991-1994 and 2001-03 [27-29]. Individuals were selected based on the Central Population Register Code to reflect the adult Danish general population aged 20-80+ years. Blood samples for DNA extraction are available for 10,594 participants. For 5,467 participants of this study ABI measurements are available from the 2001-03 examination.

1.3.3 Multi-Ethnic Study of Atherosclerosis (MESA)

MESA is an NHLBI-sponsored population-based, prospective, multi-center cohort study including participants recruited from six field sites in the United States – Forsyth County, NC (Wake Forest University), Northern Manhattan/Bronx, NY (Columbia University), Baltimore/Baltimore County, MD (Johns Hopkins University), St. Paul, MN (University of Minnesota, Twin Cities), Chicago, IL (Northwestern University), and Los Angeles County, CA (UCLA). Details of recruitment and study design have been previously published elsewhere [30]. Briefly, the MESA cohort comprises 6,814 men and women of diverse ethnic background who were 45 to 84 years old at the baseline exam and free of clinically overt cardiovascular disease (CVD) who were recruited to elucidate the determinants and natural history of subclinical CVD, and study progression of subclinical CVD. The cohort is 53% women with an ethnic composition of approximately 40% White, 30% African American, 20% Hispanic and 10% Asian primarily of Chinese descent. The MESA was approved by the Institutional Review Board of all participating field sites and reading centers and participants gave informed consent for participation and use of DNA specimens. Genotyping took place

using the Affymetrix 6.0 array. There were n=2611 participants of European descent who consented the use of DNA for research studies and had imputed GWAS data for the contributed SNPs as well as ABI measurements.

1.3.4 National Health and Nutrition Examination Surveys (NHANES)

The National Health and Nutrition Examination Surveys (NHANES) are conducted by the National Center on Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC). NHANES is a collection of diverse, population-based cross-sectional surveys of non-institutionalized Americans regardless of health status at the time of ascertainment (http://www.cdc.gov/nchs/nhanes.htm). NHANES is considered a complex survey given that specific age groups (such as the elderly) and racial/ethnic groups (non-Hispanic blacks and Mexican-Americans) are oversampled. The NHANES data accessed for this work includes NHANES 1999-2000 and NHANES 2000-2001. Collectively, these surveys contain 7,839 DNA samples linked to demographic, health, and lifestyle data. Participants were consented by the CDC at the time of the survey and sample collection, and consent included the storage of data and biological specimens such as blood for future research [31,32]. The present study was approved by the CDC Ethics Review Board. Because the study investigators did not have access to personal identifiers, this study was considered non-human subjects research by the Vanderbilt University Internal Review Board.

Race/ethnicity is self-described and was categorized as non-Hispanic white, non-Hispanic black, Mexican-American, and others. Only data for non-Hispanic whites were included here for replication efforts. The ankle brachial pressure index (ABPI), also referred to as the ankle arm blood pressure index (AAI), is a noninvasive technique to measure peripheral vascular disease in population-based and clinical studies. ABPI was defined as the ratio of the systolic blood pressure in the ankle compared with that in the arm. ABPI measurement was performed by a health technician on all adults ≥40 years of age at interview. Exclusions for these procedures were bilateral amputation, casts, ulcers, dressings, or other conditions which make blood pressure readings at these sites impossible. The presence of these conditions on one limb did not exclude the study participant, but they were excluded if they occurred bilaterally. Systolic pressure was measured in one arm (brachial vessel, right arm if accessible) and both ankles. ABI was defined as the ratio of the systolic blood pressure in the ankle compared with that in the arm. The lower ABI of the two legs was used in the analysis. ABI > 1.4 and equal to 0 were excluded from the analysis. Serum HDL-C, triglycerides, and total cholesterol were measured using standard enzymatic methods. LDL-C was calculated using the Friedewald equation, with missing values assigned for samples with triglyceride levels greater than 400 mg/dl. Fasting status was recorded at time of blood draw. Body mass index was calculated from height and weight measured in the Mobile Examination Center by CDC medical personnel. Current smoking was defined by “do you smoke cigarettes now?” or cotinine levels > 15ng/ml. Former smoking was defined among non-current smokers by the answer “yes” to the question “Smoked at least 100 cigarettes in life“. Participants were considered to have type 2 diabetes if they answered “yes” to “Ever been told you have sugar/diabetes?” or if they had fasting blood glucose levels >126 mg/dL. Hypertension was defined if average systolic blood pressure was≥140 or average diastolic blood pressure was ≥90 or if the study participant reported taking medication for high blood pressure. Prevalent cardiovascular disease was defined by “Ever told you had heart attack”. Lipid lowering medication use was defined by “now taking prescribed medicine” to lower cholesterol.

Genotyping was performed for nine SNPs in all of NHANES 1999-2002 (n=7,839) using Sequenom in the Vanderbilt DNA Resources Core. SNPs targeted for genotyping included rs10757269, rs16824978, rs4659996, rs7003385, rs7100623, rs819750, rs9485528. In addition to genotyping experimental NHANES samples, we genotyped blind duplicates provided by CDC. All SNPs reported here save for rs16824978 passed CDC quality control metrics and are available for secondary analyses through NCHS/CDC.

Tests of association were performed unweighted using linear regression, unadjusted and adjusted for covariates as specified in the analysis plan. All statistical analyses were conducted remotely in SAS v9.2 (SAS Institute, Cary, NC) using the Analytic Data Research by Email (ANDRE) portal of the CDC Research Data Center in Hyattsville, MD.

1.3.5 PREVEND

The Prevention of REnal and Vascular ENd stage Disease study is an ongoing prospective study investigating the natural course of increased levels of urinary albumin excretion and its relation to renal and cardiovascular disease. Inhabitants 28 to 75 years of age (N=85,421) in the city of Groningen, The Netherlands, were asked to complete a short questionnaire, 47% responded, and individuals were then selected with a urinary albumin concentration of at least 10 mg/L (N= 7,768) and a randomly selected control group with a urinary albumin concentration less than 10 mg/L (N=3,395). Details of the protocol have been described elsewhere (www.prevend.org). Genotyping was performed in a subset of subjects. All subjects provided written informed consent.

1.3.6 Cardiovascular Disease in Intermittent Claudication (CAVASIC) Study

The CAVASIC Study (CArdioVAScular disease in Intermittent Claudication) is a prospective case-control study which was initiated in 2002 to identify cardiovascular risk factors in patients with intermittent claudication [(33)]. Patients and controls were enrolled in two clinical centers, the Department of Vascular Surgery, Medical University Innsbruck and the 3rd Medical Department of Metabolic Diseases and Nephrology, Hietzing Hospital, Vienna, Austria. Patients were consecutively included in the study when they presented with or had a history of intermittent claudication (PAD IIa or IIb according to the criteria of Fontaine), regardless of whether they had already undergone a treatment procedure (bypass surgery or intervention). Patients were excluded from the study for any of the following: presence of acute or critical limb ischemia (Fontaine III or IV); impaired liver function with elevated enzymes (AST>50U/L, ALT >25U/L, Gamma GT >60U/L); impaired kidney function with serum creatinine >1.5 mg/dL; malignancy; past organ transplantation; therapy with nicotinic acid or corticosteroids. Furthermore controls were recruited from the same geographical region matched for age and presence of type 2 diabetes mellitus (T2DM). All members of the control group volunteered to participate in the study following a public invitation in newspapers. We applied the same exclusion criteria to the control group as used for patients. Controls with symptomatic PAD were excluded, but those with a history of cardiovascular disease were allowed to participate.

Neither the patients nor the controls suffered from acute illnesses or clinically detectable inflammatory processes at the time of enrollment. All subjects provided written informed consent and the examination protocol was approved by the Ethical Committee of the participating study centers. To minimize interobserver bias all interviews and examinations were performed by one physician at each of the two clinical centers who was specially trained in vascular examinations and echocardiography.

1.3.7 GENEPAD

The GENEPAD study was approved by the Institutional Review Boards at Stanford University and Mount Sinai; and registered at http://clinicaltrials.gov (NCT 00380185). Participants were recruited from patients undergoing elective, non-emergent coronary angiogram for angina, shortness of breath, abnormal stress test, or known CAD at Stanford University and Mount Sinai Medical Centers between April 2004 and February 2008, for whom the PAD status was not known to the investigators at the time of informed consent and recruitment into the study. This cross-sectional design allowed for two sharply defined phenotypes to emerge (i.e., patients with hemodynamically significant CAD alone, versus those with hemodynamically significant CAD and PAD). Cases were defined as patients with CAD and PAD while controls were defined as subjects with CAD alone. Furthermore, with this study design, confounding clinical covariates (e.g., traditional cardiovascular risk factors) would contribute minimally to the difference between the two phenotypes. This design would maximize our chances for discovery of novel factors that affect the distribution of atherosclerotic plaque. All participants provided written informed consent. Patients admitted for emergent catheterizations, or screening catheterizations prior to organ transplants were excluded. Additional exclusions included being younger than 40 years of age and having a history of radiation treatment, known chronic infectious diseases such as HIV, hepatitis B or C, or a language barrier.

1.3.8 Linz Peripheral Arterial Disease (LIPAD) Study

The LIPAD Study was performed from April 2000 to April 2002 at the St. John of God Hospital, Department of Surgery, Linz, Austria. Of the patients admitted for inpatient evaluation of suspected or definite PAD during the given time interval, all patients with chronic atherosclerotic occlusive disease of the lower extremities associated with typical symptoms were included into this study on the basis of the final clinical diagnosis established by the attending vascular surgeons. The diagnosis was verified by interview, physical examination, noninvasive techniques, and angiography, as detailed below. All cases with acute ischemia were excluded. Further exclusion criteria were PAD caused by nonatherosclerotic causes (cardioembolic disease, thromboangiitis obliterans, vasculitis, or congenital or metabolic vascular disease) and the history or presence of any malignancy. Doppler segmental blood pressure of the lower limbs, including continuous-wave spectral analysis and resting ankle-brachial index (ABI) measurements were performed in all patients with symptomatic PAD and control subjects as previously described [(34)]. In addition to these measurements, intra-arterial aortofemoral angiography was performed in all patients to confirm the presence of PAD and to determine the location and extent of wall changes.

Control subjects were matched to the patients with PAD in a 1:1 design by sex, age (±2 years), and diabetes mellitus status. All control subjects were patients in our hospital and fulfilled the following criteria: no clinical indication of PAD by history and physical examination; ABI ≥0.9; no pathologic pattern of pulse waves in lower limbs by continuous-wave spectral analysis; no CAD; no CVD; no previous vascular surgery or stenting of the internal carotid arteries; no stenosis of the internal carotid artery greater than 50% by color duplex ultrasound scans; no history of venous thromboembolism; and no history or presence of any malignancy. All control subjects were generally in good health and were admitted for treatment of minor health problems, such as cataract surgery, vertebragenic pain, or nonvascular surgery. The study protocol was approved by the local ethics committee in accordance with the Declaration of Helsinki, and all study participants gave informed consent.

2. Bioinformatic Analysis

Data mining in publicly available databases was used to search for biological support for the signals selected for replication. Transcript structure and functional descriptions for every gene located +/- 50 kb from each SNP (+/- 200 kb for intergenic SNPs) were automatically retrieved from different bioinformatic databases (NCBI Entrez RefSeq, EBI SRS, and the UCSC Genome Browser) using Genowatch (now named Variowatch) [(35)] and further supplemented for interesting candidates by ad-hoc searches in OMIM and Pubmed. Additionally, the Genetic Association Database, the NHGRI GWAS Catalog [(36)] and the HuGE Navigator [(37)] database were mined for already known phenotype associations. Finally, in order to refine the functions of genes without clearly associated functions in the public databases pathway, we evaluated the pathway assignments in NCBI Entrez Gene were and used Genomatix Bibliosphere Software (Genomatix GmbH, Munich, Germany) to perform an automated literature and pathway analysis. The generated gene networks were then inspected for relevant tissue types and disease associations using build-in filter functions. However, this approach was inconclusive, probably due to the little data available in PubMed, and was thus not further pursued.

3. Statistical methods used for the candidate gene look-up

In the following, the statistical methods used for the look-up of PAD/ABI and CAD candidate genes will be explained in detail, exemplified on the *CYBA* gene region:

1. All SNPs were identified, which lie in the chromosomal region defined by each of the candidate genes ±100 kb upstream or downstream of the gene. In the case of *CYBA*, 138 SNPs lie within this defined region.
2. We then identified the most strongly associated SNP in each gene region using the discovery meta-analysis results for ABI (excluding isolates): for *CYBA* this was rs3794624 with a p-value of 6.30E-05.
3. Since the lowest p-value within each gene region was selected adjusting for multiple testing within each gene region is necessary. Neighboring and densely spaced SNPs are likely to be in high LD. Furthermore, the SNPs are given as genotype scores (allelic dosages between 0 und 2) derived from an imputation algorithm using HapMap. In this situation with high correlation between SNPs, simple Bonferroni correction with the number of SNPs would be too conservative. Therefore, the effective number of loci was calculated for each gene region using an algorithm by Gao et al [(38)]. This method reduces the dimension of the parameter space. In the case of *CYBA*, the 138 SNPs could be explained sufficiently by a linear combination of 58 independent variables. Now, this so called effective number of loci can be used as the denominator in the Bonferroni correction formula. This leads to a corrected p-value of 6.30E-05 / 58 = 3.60E-03 for the *CYBA* gene region.
4. So far, we have applied the Bonferroni correction within each gene region separately resulting in gene-region-wide corrected p-values. Since we have looked at 55 independent gene regions, further correction, however, is needed. The False Discovery Rate (FDR [(39)]) was calculated on the 55 corrected p-values. This approach accounts for the multiple comparisons problem. Traditional methods like Bonferroni control the Family-wise error rate, the probability for at least on false positive among all tests. The False discovery rate on the contrary is the expected proportion of false positives amongst all rejected null hypothesis. It has a greater power than the methods controlling the Family-wise error rate

with the higher risk of false positives, however. In our analysis we have set the FDR threshold to 0.1, meaning that we are tolerating 10% of false positives out of the resulting list of significant findings. Since we are only including former PAD/ABI candidate gene regions, we are convinced that the prior knowledge outweighs the higher risk of false positives. For *CYBA*, a FDR of 0.0665 has been calculated.

5. For the CAD candidate genes, we have been looking at the specific reported SNPs rather than complete gene regions. In conclusion, we are sparing us the step of accounting for the number of SNPs within each gene region. Therefore, we are calculating the FDR directly on the p-valued derived from the discovery meta-analysis results for ABI.

Supplementary Table 1. ABI Protocol and Calculation for Discovery Cohorts and Replication Studies

Cohort (ABI measurement date)	ABI protocol	ABI calculation
Discovery Cohorts		
ARIC (1987-1989)	Systolic blood pressures were measured with Dinamap 846 Sx automated oscillometric device in one arm and one ankle (posterior tibial artery).[96] The ankle was randomly selected and measurement was taken in the supine position. Arm blood pressure was measured in the seated position and usually taken in the right arm. Repeat measures were obtained for the ankle about 5 to 8 minutes apart and for the arm about 5 minutes apart.	Mean systolic blood pressure in the ankle / mean of the seated arm systolic blood pressure
CHS (1989-1990)	Systolic blood pressure obtained with Doppler stethoscope (8 MHz, Parks Electronics, Aloha, OR) in right arm and both ankles (posterior tibial artery) and duplicate measures obtained for each location. [97]	The minimum of the mean systolic blood pressure in each ankle / mean of the right arm systolic blood pressure
FamHS (1994-1995)	Systolic blood pressures were measured with Dinamap 1846 Sx automated oscillometric device (Critikon, Inc., Tampa, FL). The ankle was randomly selected and a single systolic blood pressure was taken over the posterior tibial artery with the participant resting in the supine position. After three arm blood pressures were taken over the brachial artery, usually in the right arm, about 5 minutes apart.	Single systolic blood pressure in the ankle / mean of the three right arm systolic blood pressures.
FHS (Cohort:1994-1995; Offspring 1995-1998)	Participants rested for a minimum of five minutes in the supine position on the examining table prior to blood pressure measurement. Systolic blood pressure obtained with Doppler device (Parks Medical Electronics, Inc) in both arms and both ankles (posterior tibial artery) and repeat measurement in each limb obtained.[98,99] Dorsalis pedis pressure was measured if posterior tibial artery pressure could not be obtained.	Mean systolic blood pressure in each ankle site / higher of the two brachial systolic blood pressures
GENOA (1996)	ABI was measured in the supine position following 5 minutes of rest. Single systolic blood pressure obtained with Doppler device (Medisonics, Minneapolis MN) in both arms and both ankles (posterior tibial artery and dorsalis pedis).[100]	Systolic blood pressure in each ankle site (posterior tibial and dorsalis pedis/ higher of the two arm pressures. Lowest of the four ABI ratios was used for analysis.[20]
GHS (2007-2008)	ABI was measured in supine position after 10 minutes of rest. Systolic blood pressure was measured at the left arm with a Omron HEM 705-CP II. The peripheral systolic blood pressure was measured at the posterior tibial artery at both ankles with a hand-held 8-MHz Doppler probe (handydop, Elcat) and a conventional aneroid sphygmomanometer.	To calculate the ABI, the mean systolic blood pressure of the posterior tibial artery at both ankles was divided by the systolic blood pressure of the left arm.
Health ABC (1997-1998)	Arterial systolic blood pressure (SBP) was measured by a hand-held, 8-MHz Doppler probe (Huntleigh Technology, Inc., Manalapan, NJ, USA) placed directly over the artery and a conventional mercury sphygmomanometer. [42]	Means of the first and second SBP measurements for each leg and right arm were used to attain ABI. ABI was defined as the lowest ratio of SBP of either the right ankle to the right upper-arm or the left ankle to the right upper-arm.

InCHIANTI (1998)	Systolic blood pressure was obtained with a hand-held Doppler stethoscope (Parks Electronics model 41-A, Aloha, OR) in the right brachial and both posterior tibial arteries with duplicate measurements.[101]	The highest systolic pressure at each site was used. The ABI was calculated using the lower of the right and left posterior tibial pressures / brachial artery pressure.
KORA F3 (2004-2005)	Systolic blood pressure obtained with Doppler device (HNE Healthcare, Mini Doppler Modell Nr. D 900) in the right arm and both ankles (posterior tibial artery). We generally performed two measurements and used the mean of the two measurements for further calculations. If we observed a difference of 10 mmHg or more for the two measurements, we performed a third measurement and calculated the mean of the two measurements which were closest to each other.	Mean systolic blood pressure of that ankle side which was lowest / mean of the two brachial systolic blood pressures of the right arm
KORA F4 (2006-2008)	Systolic blood pressure obtained with Doppler device (HNE Healthcare, Mini Doppler Modell Nr. D 900) in the right arm and both ankles (posterior tibial artery). We generally performed two measurements and used the mean of the two measurements for further calculations. If we observed a difference of 10 mmHg or more for the two measurements, we performed a third measurement and calculated the mean of the two measurements which were closest to each other.	Mean systolic blood pressure of that ankle side which was lowest / mean of the two brachial systolic blood pressures of the right arm
NESDA (2004-2007)	Both ankle and arm systolic blood pressure was measured by an ultrasound Doppler device at 8-MHz (UltraTech PD1v, Ultrasound Technologies Ltd, Itton, Chepstow, UK) in combination with an ordinary blood pressure cuff, as previously described. Blood pressure was assessed with the respondent in supine position.	Ankle-brachial index was calculated as the mean of two consecutive systolic right posterior tibial artery blood pressures divided by the mean of two consecutive systolic right humeral artery blood pressures.
NBS (2005-2008)	To measure ABI appropriately sized cuffs were placed around both arms above the elbow and around both legs just above the ankle. Resting blood pressures were measured at the left and right brachial artery, the left and right posterior tibial, and dorsalis pedis arteries using an 8-MHz hand-held Doppler probe (IMEXDOPCT; Biomedic, Almere, The Netherlands).	The highest of the two arm pressures was used to calculate ABI at rest for the posterior tibial and dorsalis pedis arteries. The lowest of the four ABI's was used in the analysis.
RS-I (1990-1993) RS-II (1999-2001)	Two seated blood pressures measured in right arm and a single systolic blood pressure obtained in each leg in supine position (posterior tibial) using Doppler device (Huntleigh 500 D, Huntleigh Technology) [102] and random-zero manometer.	Systolic blood pressure in each ankle / mean of the seated right arm systolic blood pressure. The lowest ABI in either leg was used in the analysis.
SHIP (2008-2009)	Systolic blood pressure was measured with a "Dopplex D900" (Huntleigh Healthcare Ltd.) doppler ultrasound probe and a blood pressure cuff (Welch Allyn) in both arms and both ankles (anterior and posterior tibial artery). Measurements were taken in the supine position after at least 10 minutes rest.	The highest systolic pressure at each ankle site was used. The ABI was calculated using the lower of the right and left tibial pressures / higher of the two brachial artery pressures.

Discovery Cohorts Population Isolates		
Amish (2001-2009)	Systolic blood pressure was measured with a Doppler ultrasound probe (Hokanson, MD6 Bidirectional Doppler) and a blood pressure cuff (Hokanson blood pressure cuff in two sizes, SC10 and SC12)in both arms and both ankles (anterior and posterior tibial artery). Measurements were taken in the supine position after at least 15 minutes rest.	The ABI for each leg was calculated separately by dividing the ankle pressure in that leg by the mean brachial pressure of the two arms or the highest of the 2 arms if the difference between the two arms was >=10 mmHg . Then the mean ABI for both legs was used for analysis.
Croatia-Vis (2003-2004) Croatia-Korcula (2007) Croatia- Split (2009-2010)	Participant asked to remove all tight clothing from arms and legs and lie flat for approx 5 min. Brachial measure: sphymomanometer cuff placed around the right arm just above the elbow. Doppler probe placed over the brachial pulse. Cuff inflated until the Doppler sound disappears then cuff slowly deflated until the sound returns. Repeated on left arm. Ankle measure: Cuff placed around left leg just above the malleolus and Doppler used to locate the dorsalis pedis pulse. Cuff inflated until the Doppler sound disappears then cuff slowly deflated until the sound returns. Repeated for posterior tibial pulse. Both repeated for right leg.	The lowest posterior tibial systolic blood pressure at either ankle was divided by the highest systolic blood pressure in the arm
ERF (2002)	Two seated blood pressures measured in right arm and a single systolic blood pressure obtained in each leg (posterior tibial) using Doppler device (Huntleigh 500 D, Huntleigh Technology) [(102)] and random-zero manometer.	Systolic blood pressure in each ankle/ mean of the seated right arm systolic blood pressure
Orcades (2005-2007)	Participant asked to remove all tight clothing from arms and legs and lie flat for approx 5 min. Brachial measure: sphygmomanometer cuff placed around the right arm just above the elbow. Doppler probe placed over the brachial pulse. Cuff inflated a further 10-20 mm Hg after the Doppler sound disappears then cuff slowly deflated until the pulse sound returns - measure recorded. Repeated on left arm. Ankle measure: Cuff placed around left leg just above the malleolus and Doppler used to locate the dorsalis pedis pulse. Cuff inflated a further 10-20 mm Hg after the Doppler sound disappears then cuff slowly deflated until the pulse sound returns - measure recorded. Repeated for posterior tibial pulse. Both repeated for right leg.	The lowest posterior tibial systolic blood pressure at either ankle was divided by the highest systolic blood pressure in the arm
Replication Studies		
Bruneck Study (1995)	The ankle-brachial index (ABI) was measured in a supine position. A cuff was inflated to 10 mmHg above systolic blood pressure and deflated at 2 mmHg/s. The first reappearance of the arterial signal at the ankle (posterior tibial artery) was taken as the systolic blood pressure (detected with a Doppler ultrasonic instrument).	To calculate the ABI for the right and left leg, the systolic blood pressure at each ankle was divided by the systolic blood pressure in the arm. The higher arm reading (right or left side) was used for ABI calculation. The lower ABI of the two legs was considered in the current analysis. Subjects with an ABI > 1.4 were

		excluded. There was no subject with an ABI of zero on both legs.
CCHS (2001-2003)	A standard brachial systolic and diastolic blood pressure was recorded on both arms, and systolic ankle blood pressure of the posterior tibial artery on both legs was obtained by Doppler (Huntleigh Mini Dopplex Doppler D900, Huntleigh, United Kingdom).	The ABI was the lowest ankle systolic blood pressure divided by the highest brachial systolic blood pressure.
MESA (2000-2002)	Systolic blood pressure obtained with Nicolet Doppler apparatus (EN50 LE 100, Nicolet vascular, Golden, CO) in both arms and both ankles (posterior tibial and dorsalis pedis). Approximately 20 elapsed between each pressure.	ABI is the minimum of the right and left ABI. The right ABI is calculated as the (maximum of the right dorsalis pedis and right posterior tibial)/ (mean right and left brachial), and the left ABI is calculated as the (maximum of the left dorsalis pedis and left posterior tibial)/ (mean of the left and right brachial). For right ABI and left ABI, if the two brachial (arm) BPs differ by 10 mmHg or more, use the higher arm pressure as the denominator.
NHANES (1999-2002)	The ankle brachial pressure index (ABPI), also referred to as the ankle arm blood pressure index (AAI), is a noninvasive technique to measure peripheral vascular disease in population-based and clinical studies. ABPI was defined as the ratio of the systolic blood pressure in the ankle compared with that in the arm. ABPI measurement was performed by a health technician on all adults ≥40 years of age at interview. Exclusions for these procedures were bilateral amputation, casts, ulcers, dressings, or other conditions which make BP readings at these sites impossible. The presence of these conditions on one limb did not exclude the study participant, but they were excluded if they occurred bilaterally. Systolic pressure was measured in one arm (brachial vessel, right arm if accessible) and both ankles (posterior tibial vessels) for ages 40 and above.	ABI was defined as the ratio of the systolic blood pressure in the ankle compared with that in the arm. The lower ABI of the two legs was used in the analysis. ABI > 1.4 and equal to 0 were excluded from the analysis.
PREVEND (1997-1998)	At the first visit, while the participant was in a supine position, systolic blood pressure was measured each minute with an automatic Dynamap XL Model 9300 series device at the right brachial artery (ten times total). Within the first 5 min, systolic ankle pressure at each leg was measured at the same time a brachial pressure was performed. For ankle pressure measurements, the posterior tibial artery was measured using an 8-MHz continuous-wave Doppler probe (Huntleigh Model D900, Huntleigh Diagnostics) and a random-zero sphygmomanometer.	The ABI was calculated as the ratio of the systolic blood pressure of the ankle and arm for each leg. The lowest ankle-brachial index in either leg was used in the analysis.
CAVASIC (2002)	For the ABI measurement, the systolic blood pressure was measured in both arms and then two additional measures were obtained on the arm with the higher systolic blood pressure. The systolic blood pressure was obtained three times for both the posterior tibial and dorsalis pedis artery in each leg. [(33)]	The mean of the second and third systolic blood pressure measurement at each site was used to calculate the ABI for each of the 4 lower extremity sites. The ABI was calculated as the ratio of the mean systolic blood pressure at each of four sites to the mean systolic blood pressure in the arm.

		The lowest ABI from the four sites was used in analysis.
GenePAD (2004-2008)	The ABI (i.e. the ratio of ankle and arm systolic blood pressure) was measured using previously established methods.[103,104] In brief, prior to undergoing coronary angiogram, each participant rested in the supine position for 5 minutes, then, using a 5-MHz Doppler ultrasound (Nicolet Elite 5-MHz vascular model 110R Doppler; Nicolet Vascular, Golden, CO, USA), systolic pressures were measured in the posterior tibial, dorsalis pedis, and brachial arteries. Each pressure was measured twice in sequential and reverse order as listed. With an IV in one arm, only the contralateral arm was used for the brachial pressures. [105]	The ABI for each leg was calculated separately by dividing the higher of the two ankle pressures in that leg by the brachial pressure. If the ABI was less than 0.9 in either leg, the patient was considered to have PAD. The index leg was defined as the leg with the lower ABI.
LIPAD (2002)	ABI measurements in the LIPAD study were done according to [34]: with the patient placed in a supine position, the brachial and ankle systolic pressure measurements are obtained.	The higher systolic pressure of the anterior tibial or posterior tibial measurement for each foot / the higher of the left and right brachial systolic blood pressure

Supplementary Table 2. Genotyping and Imputation Information for Discovery and Replication Samples (Part 1/3)

	ARIC	CHS	Family HS	FHS	GENOA	GHS	Health ABC	InCHIANTI
Genotyping platform	Affymetrix Genome-Wide Human SNP Array 6.0	Illumina 370CNV Duo® BeadChip	Illumina 550 array, Illumina 610 array, and Illumina 1M array	Affymetrix 250K Nsp and 250K Sty mapping arrays and the Affymetrix 50K supplemental gene-focused array	Affymetrix® Genome-Wide Human SNP Array 6.0	Affymetrix Genome-Wide Human SNP Array 6.0	Illumina Human1M-Duo Array	Illumina Infinium HumanHap550 array
Sample exclusions	1) Discordant with previous genotype data 2) sex mismatch 3) first-degree relative of an included individual based on genotype data; 4) genetic outlier as assessed by Identity by state (IBS); 5) > 8 SD along any of the first 10 principal components in EIGENSTRAT	1) call rate<95%; 2) sex mismatch; 3) other sample failure	1) sex mismatch; 2) outliers identified by the IBS clustering analysis	1) call rate <97%; 2) per subject heterozygosity±5 SDs from the mean; 3) per subject large Mendelian error rate	1) call rate <95%; 2) sex mismatch	1) call rate <97%; 2) per subject heterozygosity±3 SDs from the mean	1) sample failure, 2) genotypic sex mismatch 3) first-degree relative	1) genotype call rate<98%; 2) sex mismatch
SNP exclusions	1) call rate <90%; 2) MAF<1%; 3) HWE p<10-6; 4) SNPs without chromosomal location; 5) monomorphic SNPs	1) call rate<0.97; 2) >2 replicate errors or Mendelian inconsistencies among reference CEPH trios; 3) HWE p<10-5 4) no observed heterozygotes	1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <95%; 4) rsqhat<0.3	1) call rate < 97%; 2) Mishap p< 10-9; 3) HWE p < 10-6; 4) Mendelian errors > 100	1) call rate <95%; 2) monomor-phic SNPs; 3) MAF<1%; 4) unknown chromosomal location	1) call rate < 98% 2) MAF < 1% 3) HWE $p<10^{-4}$		1) MAF <1%; 2) HWE $p<10^{-4}$; 3) SNP call rate <98%
Imputation method	MACH (version 1.0.16)	BIMBAM version 0.99	MACH 1.0	MACH (version 1.0.15)	MACH 1.0.16	IMPUTE v. 1.0.0	MACH 1.0.16	IMPUTE 3.1
Imputation backbone	BRLMM to Hapmap-V1 CEU from build 35	HapMap CEU using release 22, build 36	HapMap release 22 CEU	HapMap (release 22, build 36, CEU population	HapMap release 22, build 36, CEU	HapMap (rel. 22, build 36, CEU pop.	HapMap II (rel. 22, build 36) CEU	HapMap II (rel. 22, build 36) CEU
Imputation SNP exclusions	SNPs with <95% data completeness, MAF <1%, HWE P<0.00005		1) MAF <1%; 2) HWE p<10-6; 3) SNP not in HapMap	1) MAF<0.01; 2) HWE p<10-6; 3) callrate >0.97; 4) mishap test of non-random missingness p>10-9; 5) <100 Mendelian errors			MAF<0.01; HWE P<10-6; callrate<97%	
Statistical software	PLINK, Mach2QTL	R	SAS proc mixed and GEE	R packages kinship, GEE	R packages nlme, MASS	SNPTEST	R	SNPTEST

Supplementary Table 2. Genotyping and Imputation Information for Discovery and Replication Samples (Part 2/3)

	KORA F3	KORA F4	NESDA	NBS	RS-I	RS-II	SHIP
Genotyping platform	Affymetrix GeneChip Human Mapping 500K	Affymetrix Genome-Wide Human SNP Array 6.0	Perlegen	Illumina HumanHapCNV370-Duo-Beadchip	Illumina 550K array	Illumina 550K array	Affymetrix Genome-Wide Human SNP Array 6.0
Sample exclusions	1) sample call rate<93%; 2) sex mismatch	1) sample call rate<93%; 2) sex mismatch	call rate<95%, ethnic outliers, XO and XXY samples, high genome-wide homo- or heterozygosity, excess IBS	1) call rate < 96%; 2) <90% Caucasian ancestry as identified by Structure analysis	1) excess autosomal heterozygosity; 2) sex mismatch; 3) outliers identified by the IBS clustering analysis	1) excess autosomal heterozygosity; 2) sex mismatch; 3) outliers identified by the IBS clustering analysis	1) call rate<92%; 2) sex mismatch;
SNP exclusions			callrate <95%, MAF<1%		1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <98%	1) MAF <1%; 2) HWE p<10-5; 3) SNP call rate <90%	NA
Imputation method	Mach 1.0.10	Mach 1.0.16	Impute 0.4.2	Impute 0.5.0	MACH 1.0	MACH 1.0.15	IMPUTE 0.5.0
Imputation backbone	HapMap (release 21, build 35, CEU)	HapMap (release 22, build 36, CEU)	HapMap Phase II - release 22 build 36, CEU	HapMap Phase II - release 22 build 36, CEU	HapMap release 22 CEU	HapMap release 22 CEU	HapMap release 22 CEU
Imputation SNP exclusions	MAF <5%	genotype callrate<93%	MAF<0.01; HWE P<10-6; callrate<95%	1) call rate <96% 2) MAF <1% 3) HWE p<10^{-6}	1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <98%	1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <98%	NA
Statistical software	Mach2QTL	Mach2QTL	SNPTEST v1.1.4; R	SNPTEST v 1.1.5	ProbABEL (version 1.1) *	ProbABEL (version 1.1) *	QUICKTEST version 0.95

* (http://mga.bionet.nsc.ru/~yurii/ABEL/).

Supplementary Methods Table 2. Genotyping and Imputation Information for Discovery and Replication Samples (Part 3/3)

	Population Isolates						Replication Studies*	
	Amish	Croatia-Vis	Croatia-Korcula	Croatia- Split	ERF	Orcades	MESA	PREVEND
Genotyping platform	Affymetrix GeneChip Human Mapping 500K	Illumina HumanHap300v1	Illumina 370CNV	Illumina 370CNV	Illumina 6K (n=3000), 318K (n=1200), 370K (n=100), Affymetrix 250K (n=200)	Illumina HumanHap300v2	Affymetrix 6.0 1M	Illumina CytoSNP12 V2
Sample exclusions	1) sample call rate<93%; 2) sex mismatch 3) excess Mendelian error rate	discrepancies between sex in phenotype file and genotyping; too high autosomal heterozygosity (FDR<1%); too high IBS sharing (>=95%); call rate (<97%)	discrepancies between sex in phenotype file and genotyping; too high autosomal heterozygosity (FDR<1%); too high IBS sharing (>=95%); call rate (<97%)	discrepancies between sex in phenotype file and genotyping; too high autosomal heterozygosity (FDR<1%); too high IBS sharing (>=95%); call rate (<97%)	1) excess autosomal heterozygosity; 2) sex mismatch; 3) outliers identified by the IBS clustering analysis	discrepancies between sex in phenotype file and genotyping; too high autosomal heterozygosity (FDR<1%); too high IBS sharing (>=95%); call rate (<97%)	1) gender mismatches and 2) cryptic duplicates	3SD phenotypic outliers, duplicates, contaminated, relatedness, gender mismat
SNP exclusions	1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <95%	MAF<0.02, Call Rate 98%, pHWE < 1E-10	MAF<0.02, Call Rate 98%, pHWE <1E-10	MAF<0.02, Call Rate 98%, pHWE <1E-10	1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <95%	MAF<0.02, Call Rate 98%, pHWE <1E-10	None required for 7 SNPs requested	SNP call rate <95%; MAF < 0.01; HWE p< 1 x 10-5
Imputation method	MACH v-1.0.16	MACH (1.0.16)	MACH (1.0.16)	MACH (1.0.16)	MACH 1.0	MACH (1.0.16)	IMPUTE V2.1.0	BEAGLE v3.1.0
Imputation backbone	HapMap (release 22, build 36, CEU population	HapMap r22 CEU	HapMap r22 CEU	HapMap r22 CEU	HapMap release 22 CEU	HapMap r22 CEU	HapMap Phase I and II - CEU, rel. 24 - Build 36 (dbSNP b126)	Hapmap rel 23a CEU
Imputation SNP exclusions	1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <95%	MAF<0.01, Call Rate 98%, pHWE < 1E-06	MAF<0.01, Call Rate 98%, pHWE < 1E-06	MAF<0.01, Call Rate 98%, pHWE < 1E-06	MAF<0.01, Call Rate 98%, pHWE < 1E-06	None required for 7 SNPs requested		1) MAF <1%; 2) HWE p<10-6; 3) SNP call rate <95%
Statistical software	MMAP	GenABEL/ProbABEL	GenABEL/ProbABEL	GenABEL/ProbABEL	PROBabel: with mmscore	GenABEL/ProbABEL	SNPTEST	PLINK v1.07

*Genotyping of the replication samples from the Bruneck Study, Copenhagen City Heart Study, CAVASIC Study and LIPAD Study was performed using 5' nuclease allelic discrimination (Taqman) assays. Genotyping of the replication samples from NHANES was performed using Sequenom.

Supplementary Table 3: Relative excess heterozygosity analysis of SNPs in replication studies

SNP	lnREH_Fixed	P_Fixed	CI_Fixed	lnREH_Random	P_Random	CI_Random	I2	I2_CI
rs10757269	0.0138	0.4582	[-0.02, 0.05]	-0.0027	0.9281	[-0.06, 0.06]	0.4675	[0.00, 0.79]
rs4659996	-0.0154	0.4073	[-0.05, 0.02]	-0.0297	0.2695	[-0.08, 0.02]	0.3645	[0.00, 0.75]
rs7003385	0.0049	0.8142	[-0.04, 0.05]	0.0049	0.8142	[-0.04, 0.05]	0.0000	[0.00, 0.07]
rs9485528	0.0184	0.5797	[-0.05, 0.08]	0.0184	0.5797	[-0.05, 0.08]	0.0000	[0.00, 0.71]
rs819750	-0.0262	0.5384	[-0.11, 0.06]	-0.0262	0.5384	[-0.11, 0.06]	0.0000	[0.00, 0.48]
rs16824978	0.0018	0.9443	[-0.05, 0.05]	0.0018	0.9443	[-0.05, 0.05]	0.0000	[0.00, 0.68]
rs7100623	-0.0952	0.0014	[-0.15, -0.04]	-0.0773	0.1140	[-0.17, 0.02]	0.5305	[0.00, 0.81]

Supplementary Table 4. Characteristics of Discovery Cohort Participants at time of Ankle-brachial Index Measurement (Part 1/3)

Characteristic, Mean (SD) or N (%)	ARIC	CHS	FamHS	FHS	GENOA	GHS	Health ABC	InCHIANTI
N	7630	3193	1736	3572	991	3122	1562	1130
Years ABI obtained	1987-1989	1989 - 1990	1994 - 1995	1994-1998	2001-2004	2007-2008	1997-1998	1998-2000
Age, years, mean (SD)	54.3 (5.7)	72.3 (5.4)	50.0 (12.9)	61.6 (11.8)	58.6 (10.1)	55.8 (10.9)	73.8 (2.8)	67.6 (15.3)
Women, n (%)	4073(53.4)	1950 (61.1)	941 (54.2)	1950 (54.6)	564 (56.9)	1533 (49.1)	743 (47.6)	613 (55.0)
ABI, mean (SD)	1.12 (0.13)	1.08 (0.14)	1.19 (0.12)	1.1 (0.13)	1.1 (0.14)	1.02 (0.12)	1.09 (0.16)	1.04 (0.15)
PAD = ABI$\leq$0.9, n (%)	287 (3.8)	264 (8.3)	41 (2.4)	194 (5.4)	66 (6.7)	405 (13.0)	148 (9.5)	116 (10.4)
ABI = 0.91 to 1.10, n (%)	2886 (37.8)	1345 (42.1)	309 (18.7)	1214 (34.0)	387 (39.1)	1941 (62.2)	593 (42.0)	670 (60.1)
Hypertension, n (%)	1693 (22.3)	1011 (31.7)	400 (23.0)	1607 (45.0)	717 (72.4)	1640 (52.5)	727 (46.8)	338 (30.8)
Diabetes, n (%)	646 (8.5)	370 (11.7)	89 (5.2)	334 (9.4)	148 (14.9)	220 (7.0)	163 (10.5)	73 (6.7)
Current smoker, n (%)	1924 (25.2)	361 (11.3)	223 (13.2)	507 (14.3)	103 (10.4)	574 (18.4)	101 (6.5)	214 (19.2)
Former smoker, n (%)	2700 (35.4)	1306 (40.9)	734 (42.5)	1683 (47.5)	389 (39.3)	1096 (35.2)	796 (51.0)	282 (25.2)
Total cholesterol, mg/dL	214.9 (40.7)	213.2 (38.9)	202.4 (40.0)	206.0 (39.3)	197.1 (33.8)	224.5 (41.3)	201.4 (37.6)	216.0 (39.9)
HDL cholesterol, mg/dL	50.7 (16.9)	55.4 (15.9)	50.1 (14.6)	51.0 (16.2)	51.8 (15.1)	57.0 (16.0)	52.1 (16.4)	56.1 (15.1)
Triglyceride, mg/dL	137.2 (91.7)	139.9 (75.7)	148.7 (102.8)	141.7 (135.4)	157.0 (98.1)	125.6 (73.9)	152.8 (86.8)	126.3 (78.6)
BMI, kg/m^2	27.0 (4.8)	26.3 (4.4)	27.3 (5.1)	27.7 (5.0)	30.7 (6.2)	27.1 (4.7)	26.6 (4.1)	27.1 (4.1)
Claudication, n (%)	63 (0.8)	37 (1.2)	NA	131 (3.7)	NA	249 (8.0)	293 (18.9)	87 (7.9)
Prevalent CVD, n (%)	381 (5.1)	0	155 (8.9)	217 (6.1)	86 (8.7)	214 (6.9)	404 (26.3)	63 (5.7)

NA, not available

Supplementary Table 4. Characteristics of Discovery Cohort Participants at time of Ankle-brachial Index Measurement (Part 2/3)

Characteristic, Mean (SD) or N (%)	KORA F3	KORA F4	NESDA	NBS	RS-I	RS-II	SHIP
N	1581	1407	1612	544	5169	1642	543
Years ABI obtained	2004-2005	2006-2008	2004-2007	2005-2008	1991-1993	2000-2001	2008-2009
Age, years, mean (SD)	62.2 (10.1)	63.4 (7.2)	41.8 (12.4)	62.8 (5.4)	69.1 (9.0)	64.3 (7.7)	58.6 (13.0)
Women, n (%)	821 (51.9)	730 (51.9)	1112 (69.0)	279 51.3	3091 (59.8)	885 (53.9)	280 (51.6)
ABI, mean (SD)	1.12 (0.13)	1.14 (0.14)	1.14 (0.12)	1.09 (0.10)	1.05 (0.21)	1.06 (0.14)	1.12 (0.13)
PAD = ABI≤0.9, n (%)	81 (5.1)	75 (5.3)	36 (2.2)	17 (3.13)	959 (18.6)	176 (10.7)	18 (3.32)
ABI = 0.91 to 1.10, n (%)	621 (39.3)	397 (28.2)	581 (36.9)	295 (56.0)	1786 (34.6)	808 (49.2)	225 (41.4)
Hypertension, n (%)	950 (60.4)	696 (49.6)	614 (38.1)	220 (40.4)	2912 (56.3)	935 (56.9)	299 (55.1)
Diabetes, n (%)	166 (10.5)	140 (10.0)	45 (2.8)	42 (7.7)	528 (10.2)	166 (10.1)	57 (10.5)
Current smoker, n (%)	214 (13.6)	195 (13.9)	670 (41.6)	83 (15.3)	1150 (22.2)	330 (20.1)	77 (14.2)
Former smoker, n (%)	569 (36.0)	583 (41.5)	512 (31.8)	273 (50.3)	2118 (41.0)	820 (49.9)	243 (44.8)
Total cholesterol, mg/dL	221.5 (40.4)	224.1 (39.6)	198.9 (41.4)	228.2 (38.7)	255.5 (47.0)	224.7 (37.9)	212.3 (42.2)
HDL cholesterol, mg/dL	58.2 (17.0)	56.5 (14.4)	63.1 (16.9)	54.5 (14.7)	51.9 (14.3)	53.4 (14.3)	54.5 (14.1)
Triglyceride, mg/dL	171.2 (121.9)	135.6 (96.1)	114.6 (77.6)	127.5 (67.3)	NA	NA	161.2 (111.7)
BMI, kg/m^2	28.1 (4.5)	28.5 (4.8)	25.5 (4.9)	26.9 (3.9)	26.3 (3.7)	27.2 (4.2)	28.16 (4.4)
Claudication, n (%)	63 (3.99)	56 (3.98)	333 (20.7)*	N/A	88 (1.7)	19 (1.2%)	50 (9.2)
Prevalent CVD, n (%)	173 (10.94)	163 (11.58)	50 (3.1)	19 (3.5)	833 (16.1)	149 (9.1%)	75 (13.8)

NA, not available

* Question asked was: "Do you have pain in one or both calves during walking?"

Supplementary Table 4. Characteristics of Discovery Cohort Participants from population isolates at time of Ankle-brachial Index Measurement (Part 3/3)

Characteristic, Mean (SD) or N (%)	Amish	ERF	CROATIA-Vis	CROATIA-Korcula	Orcades	CROATIA-Split
N	1183	2133	897	851	693	499
Years ABI obtained	2001-2009	2002	2003-2004	2007	2005-2007	2009-2010
Age, years, mean (SD)	49.0 (16.5)	49.7 (15.2)	56.1 (15.4)	55.93 (13.33)	53.74(15.31)	49.04 (14.65)
Women, n (%)	577 (48.8)	1309 (56.9)	518 (57.7)	542 (63.7)	372(53.7)	286 (57.3)
ABI, mean (SD)	1.09 (0.11)	1.05 (0.13)	1.10 (0.12)	1.02 (0.14)	1.03 (0.11)	1.03 (0.10)
PAD = ABI<0.9, n (%)	54 (4.6)	180 (8.5)	37 (4.1)	146 (17.2)	59 (8.5)	50 (10)
ABI = 0.91 to 1.10, n (%)	553 (46.8)	1195 (56.0)	464 (51.7)	494 (58.0)	468 (67.5)	346 (69.3)
Hypertension, n (%)	153 (12.9)	637 (29.0)	196 (21.9)	328 (38.5)	153 (22.1)	122 (24.4)
Diabetes, n (%)	19 (1.6)	108 (5.0)	55 (6.1)	83 (9.8)	15 (2.2)	16 (3.2)
Current smoker, n (%)	111 (9.7)	869 (39.8)	212 (23.6)	179 (21.0)	60 (8.7)	126 (25.3)
Former smoker, n (%)	162 (14.2)	653 (29.9)	229 (25.5)	224 (26.3)	233 (33.6)	138 (27.7)
Total cholesterol, mg/dL	211 (47.7)	214.5 (42.7)	197.2 (37.9)	229.3 (47.6)	224.3 (45.2)	227.4 (48.8)
HDL cholesterol, mg/dL	55.6 (14.6)	49.4 (14.1)	42.9 (6.2)	56.5 (13.5)	64.6 (15.5)	53.4 (12.9)
Triglyceride, mg/dL	75.6 (47.3)	119.3 (68.9)	150.6 (87.7)	126.7 (73.5)	117.8 (58.5)	133.5 (93.1)
BMI, kg/m2	27.0 (4.7)	26.9 (4.8)	27.3 (4.3)	28.0 (4.1)	27.8 (4.9)	26.9 (4.2)
Claudication, n (%)	NA	43 (2.0)	NA	NA	NA	NA
Prevalent CVD, n (%)	50 (4.4)	48 (2.1)	27 (3.0)	11 (1.3)	52 (7.5)	4 (0.8)

NA, not available

Supplementary Table 5. Characteristics of the Replication Samples at the time of Ankle Brachial Index Measurement

Characteristic, Mean (SD) or N (%)	Bruneck Study	CCHS	NHANES	MESA	PREVEND	CAVASIC cases	CAVASIC controls	GenePAD	LIPAD cases	LIPAD controls
N	786	5330	2375	2611	3691	136	307	850	279	392
Years ABI obtained	1995	2001-2003	1999-2002	2000-2002	1997-1998	2002-2006	2002-2006	2004-2008	2000-2002	2000-2002
Age, years, mean (SD)	62.4 (11.0)	59.1 (16.0)	61.95(13.53)	62.7 (10.2)	49.4 (12.5)	59 (7)	57 (10)	68 (11)	68 (11)	67 (11)
Women, n (%)	397 (50.5)	57.5	1128(47.49)	1360 (52.1)	1791 (48.5)	0 (0)	0 (0)	258 (30)	87 (31)	121 (31)
ABI, mean (SD)	1.10 (0.15)	1.01 (0.15)	1.09(0.14)	1.11 (0.12)	1.16 (0.14)	0.7 (0.3)	1.1 (0.1)	1.0 (0.2)	0.7 (0.2)	1.2 (0.1)
PAD = ABI<0.9, n (%)	79 (10.1)	959 (18.0)	202 (8.5%)	82 (3.1)	135 (3.7)	106 (78)	19 (6)	153 (18)	239 (86)	0 (0)
ABI = 0.91 to 1.10, n (%)	321 (40.8)	2974 (55.8)	917(38.6)	926 (35.5)	802 (21.7)	15 (11)	161 (52)		22 (8)	121 (31)
Hypertension, n (%)	535 (68.1)	2703 (50.7)	1162(49.6%)	1000 (38.3)	1232 (33.4)	119 (88)	195 (64)	445 (52)	165 (59)	159 (41)
Diabetes, n (%)	80 (10.2)	327 (6.1)	188(8.01%)	116 (4.4)	139 (3.8)	15 (11)	97 (32)	125 (15)	89 (32)	105 (27)
Current smoker, n (%)	158 (20.1)	1736 (32.7)	529(22.3%)	296 (11.3)	1335 (36.1)	73 (55)	39 (13)	NA	125 (45)	51 (13)
Former smoker, n (%)	203 (25.8)	1847 (51.9)	908(38.2%)	1156 (44.3)	1327 (36.0)	54 (40)	140 (46)	NA	20 (7)	24 (6)
Ever smoker, n (%)	361 (45.9)		1361(57.3%)	1451 (55.6)	NA	127 (95)	179 (58)	511 (60)	145 (52)	75 (19)
Total cholesterol, mg/dL	223.2 (44.6)	212.9 (45.3)	208.71(39.18)	195.7 (35.5)	217.6 (42.5)	208 (43)	205 (35)	117 (40)	232 (52)	214 (44)
HDL cholesterol, mg/dL	56.6 (13.7)	57.1 (18.6)	52.44(16.35)	52.4 (15.7)	50.3 (15.2)	50 (13)	59 (17)	40 (12)	52 (16)	54 (17)
Triglyceride, mg/dL	135.6 (90.1)	137.9 (118.9)	162.95(154.84)	132.3 (89.1)	131.2 (89.5)	165 (98)	131 (74)	108 (68)	163 (102)	134 (76)
Lipid lowering meds, n (%)	24 (3.1)	296 (5.6)	415(17.5%)	488 (18.7)	(4.1)	51 (38)	60 (20)	689 (81)	56 (20)	0 (0)
BMI, kg/m2	25.6 (3.8)	25.8 (4.3)	27.94(5.45)	27.8 ± 5.1	26.1 (4.2)	27 (4)	27 (4)	29 (6)	26 (4)	27 (4)
Claudication, n (%)	35 (4.5)	NA	NA	10 (0.38)	NA	136 (100)	0 (0)	85 (10)	236 (85)	0 (0)
Prevalent CVD, n (%)	109 (13.9)	702 (13.1)	169(7.2)	0 (0)	143 (3.9)	39 (29)	29 (9)	520 (61)	116 (42)	0 (0)

NA, not available

Supplementary Table 6: Overall Meta-analysis Results and Study Specific Results: ABI-SNP Associations with $p<10^{-5}$, population isolates excluded

Table is at the end of this document

Supplementary Table 7. Meta-analysis results: ABI-SNP associations with $p<10^{-5}$ in the primary discovery analysis with population isolates included.

SNP	Chr	Physical Position	Closest Gene	Risk / Non-risk Allele	Risk Allele frequency	Meta-analysis	N	Beta	SE	P value	P_{het}
rs10757269	9	22062264	*CDKN2B*	G/A	0.50	ABI Discovery	41292	-0.0048	0.0009	1.22E-07	0.01
						ABI Replication	16672	-0.0035	0.0015	1.76E-02	0.67
						ABI Combined	**57964**	**-0.0045**	**0.0008**	**8.82E-09**	**0.04**
						PAD† Discovery	39600	0.0686	0.0274	1.23E-02	0.14
rs819750	1	99469651	*LPPR4*	G/T	0.12	ABI Discovery	41534	-0.0065	0.0014	2.79E-06	0.5
						ABI Replication	16660	0.0022	0.0023	3.22E-01	0.99
						ABI Combined	**58194**	**-0.0041**	**0.0012**	**5.11E-04**	**0.36**
						PAD Discovery	39825	0.0766	0.0407	5.97E-02	0.04
rs4659996	1	238912747	*GREM2*	A/G	0.48	ABI Discovery	33160	-0.005	0.0011	3.48E-06	0.34
						ABI Replication	16658	-0.0018	0.0016	2.67E-01	0.65
						ABI Combined	**49818**	**-0.004**	**0.0009**	**7.27E-06**	**0.4**
						PAD Discovery	32619	0.0467	0.0295	1.14E-01	0.33
rs11580768*	1	2789340	*ACTRT2*	C/T	0.53	**ABI Discovery**	**34252**	**-0.0053**	**0.0012**	**3.85E-06**	**0.99**
						PAD Discovery	33735	0.0832	0.0314	0.008095	0.4367
rs4366776*	17	216763	*RPH3AL*‡	C/T	0.51	**ABI Discovery**	**37127**	**-0.0047**	**0.001**	**4.52E-06**	**0.78**
						PAD Discovery	35398	0.061	0.0304	0.04508	0.4154
rs10509665*	10	95518432	*LGI1*‡	A/G	0.61	**ABI Discovery**	**41665**	**-0.0042**	**0.0009**	**5.99E-06**	**0.33**
						PAD Discovery	39958	0.0454	0.0281	0.1058	0.1607
rs10507044*	12	93774126	*NDUFA12*	T/C	0.41	**ABI Discovery**	**41596**	**-0.0041**	**0.0009**	**7.25E-06**	**0.91**
						PAD Discovery	39910	0.0386	0.0277	0.1629	0.3781

P_{het}= p value for heterogeneity

* Not chosen for replication

‡ SNP is located within the gene; rs819750 and rs4366776 are within 60kb of the gene †PAD defined as ABI≤0.9 vs ABI >0.9

Supplementary Table 8. Meta-analysis results: SNP associations for PAD (ABI ≤0.9 vs ABI >0.9) with $p<10^{-5}$ with population isolates† included.

SNP	Chr	Physical Position	Closest Gene	Risk / Non-risk Allele	Risk Allele frequency	N	OR	95% Confidence Interval	P value	P_{het}
rs9998941	4	162544312	*FSTL5**	A/G	0.22	39715	1.18	(1.11, 1.27)	6.00E-07	0.49
rs11751656	6	42751046	*UBR2**	G/A	0.07	29586	1.61	(1.33, 1.97)	1.72E-06	0.81
rs12772949	10	82698704	*SH2D4B*	G/A	0.17	38786	1.21	(1.12, 1.31)	2.14E-06	0.94
rs6426183	1	245079347	*AHCTF1**	C/T	0.09	39959	1.24	(1.13, 1.36)	4.60E-06	0.54
rs11715199	3	178757173	*TBL1XR1*	G/C	0.28	39960	1.15	(1.08, 1.22)	4.89E-06	0.43
rs7696128	4	18697484	*LCORL*	G/T	0.52	39960	1.13	(1.07, 1.20)	5.02E-06	0.88
rs3119311	6	160613097	*SLC22A2*	C/T	0.12	39958	1.2	(1.11, 1.30)	5.54E-06	0.47

P_{het}= p value for heterogeneity

* SNP is located within the gene

† The Amish Study did not participate in the meta-analysis

Supplementary Table 9. SNP associations for both ABI and PAD (ABI<0.9) with $p<10^{-4}$, excluding population isolates.

SNP	Chr	Closest Gene	Risk Allele	Non-risk Allele	Risk Alle Frequency	Ankle-brachial index (ABI)			Peripheral arterial disease (PAD)		
						Beta	P value	N	Beta	P value	N
rs615552	9	*CDKN2B*	t	c	0.5665	-0.0053	1.52E-07	35397	0.1176	8.69E-05	34903
rs523096	9	*CDKN2B*	a	g	0.5615	-0.0052	3.36E-07	35401	0.1165	9.99E-05	34915
rs518394	9	*CDKN2B*	g	c	0.5989	-0.0047	9.38E-06	35252	0.1256	8.60E-05	34904
rs7100623	10	*IDE* *	t	c	0.1992	-0.0054	1.89E-05	35412	0.1394	8.39E-05	34914
rs7908111	10	*IDE* *	a	g	0.2011	-0.0052	3.22E-05	35377	0.1384	7.90E-05	34879
rs13290547	9	*DAB2IP* *	t	c	0.0625	-0.0093	3.62E-05	32135	0.2797	2.20E-05	31720
rs7898493	10	*IDE* *	c	g	0.2032	-0.0051	3.67E-05	35346	0.1414	5.78E-05	34852
rs11187074	10	*IDE* *	c	g	0.2001	-0.0058	5.64E-05	27783	0.1380	8.77E-05	34915
rs1855916	10	*IDE* *	a	g	0.2022	-0.0050	5.87E-05	35320	0.1491	2.60E-05	34828
rs7084090	10	*IDE* *	a	t	0.2022	-0.0050	5.93E-05	35321	0.1490	2.62E-05	34829
rs6438172	3	*GRAMD1C* *	a	g	0.2165	-0.0048	6.56E-05	35413	0.1359	8.69E-05	34915
rs2245830	3	*GRAMD1C* *	g	t	0.2163	-0.0048	6.62E-05	35400	0.1364	8.83E-05	34902
rs4682143	3	*GRAMD1C* *	t	g	0.2162	-0.0048	6.64E-05	35402	0.1360	8.78E-05	34904
rs7093418	10	*IDE* *	t	g	0.2023	-0.0050	6.69E-05	35320	0.1497	2.43E-05	34828
rs2566973	3	*GRAMD1C* *	t	c	0.2161	-0.0048	6.99E-05	35400	0.1368	8.10E-05	34904
rs2566977	3	*GRAMD1C* *	c	a	0.2158	-0.0048	9.04E-05	35398	0.1366	9.90E-05	34901
rs1887922	10	*IDE* *	c	t	0.2005	-0.0049	9.74E-05	35320	0.1467	4.27E-05	34827

SNPs ordered by strength of the ABI association p value

* SNP is located within the gene

Supplementary Table 10: Literature-reported candidate genes for ankle-brachial-index (ABI) and/or peripheral arterial disease (PAD) and their association with ABI in the CHARGE GWAS discovery sample (population isolates excluded). The table reports the most strongly associated SNP (according to p-value) within the gene region ± 100 kb upstream and downstream of the candidate gene. We selected a candidate gene only if an association study with at least 100 cases and 100 controls was available independent whether the study was positive or negative. This decision was made on the basis that most of the negative studies were largely underpowered to exclude an association.

						Most strongly associated SNP within the gene region ± 100 kb upstream and downstream of the candidate gene									
Candidate gene	Investigated markers as reported	Ref.	Chr.	Coding strand	Gene region 5' start 3' end	Top hit rs number	Position (bp)	Nearest gene	Ref./ coded allele	Coded allele freq.	Effect	P-value	# of loci (eff. loci)*	corr. P-value**	False discovery rate‡
ACE	rs4340 (I/D), rs4291	(40-45)	17	(+)	58,808,166 59,052,935	rs4459609	58,902,680	ACE	C/A	0.650	-0.002	0.0907	87 (31)	1.0000	1.0000
ADD1	G460W	(46)	4	(+)	2,715,382 3,001,585	rs11947904	2,774,570	SH3BP2	G/T	0.052	0.007	0.0150	183 (63)	0.9463	1.0000
ADRB2	Two polymorphisms	(47)	5	(+)	148,086,349 148,288,381	rs1432628	148,231,861	ADRB2	C/T	0.054	-0.007	0.0012	249 (70)	0.0860	0.4298
AGT	G-6A, T174M, M235T, rs699	(41,42, 44)	1	(-)	228,804,892 229,016,564	rs4847000	228,848,364	COG2	G/A	0.219	0.003	0.0369	373 (109)	1.0000	1.0000
AGTR1	573C/T (exon 5), rs5186	(41,44)	3	(+)	149,798,348 150,043,480	rs1492094	150,028,691	CPB1	G/A	0.392	0.003	0.0151	334 (88)	1.0000	1.0000
BRAP	rs11066001	(48)	12	(-)	110,466,279 110,708,122	rs2051792	110,596,100	BRAP	C/T	0.024	0.008	0.0399	55 (18)	0.7173	1.0000
CDKN2B	rs1333049	(49)	9		22,015,503 22,215,503	rs10757269	22,062,264	CDKN2B	G/A	0.506	0.006	2.5E-08	213 (69)	1.7E-06	9.32e-05
Chr 11 region***	rs9665943, rs1042602	(50)	11		82,027,366 88,651,344	rs4144290	84,208,186	DLG2	G/A	0.126	-0.005	0.0058	7014 (1577)	1.0000	1.0000
CHRNA3	rs1051730	(51,52)	15	(-)	76,574,707 76,800,377	rs1878399	76,699,058	CHRNA3	G/C	0.575	-0.003	0.0047	191 (40)	0.1869	0.5943
CRP	1059 G/C	(53)	1	(-)	157,848,704 158,051,003	rs2808627	157,935,964	CRP	T/A	0.237	0.003	0.0197	221 (74)	1.0000	1.0000
GJA4 (CX37)	1019C>T Pro319Ser rs1764391	(54,55)	1	(+)	34,931,186 35,133,933	rs12127690	35,077,856	C1orf212	G/C	0.689	0.003	0.0229	124 (40)	0.9176	1.0000
CX3CR1	V249I, T280M	(56)	3	(-)	39,179,990 39,396,531	rs2669845	39,296,222	CX3CR1	C/T	0.133	-0.004	0.0073	192 (65)	0.4767	1.0000

to be continued

Supplementary Table 10: continued

						Most strongly associated SNP within the gene region ± 100 kb upstream and downstream of the candidate gene									
Candidate gene	Investigated markers as reported	Ref.	Chr.	Coding strand	Gene region 5' start 3' end	Top hit rs number	Position (bp)	Nearest gene	Ref./ coded allele	Coded allele freq.	Effect	P-value	# of loci (eff. loci)*	corr. P-value**	False discovery rate‡
CYBA (p22 phox)	C242T	(57)	16	(-)	87,137,199 87,344,958	rs3794624	87,244,575	CYBA	G/A	0.340	0.005	6.3E-05	138 (58)	0.0036	0.0665
CYP2C9	rs1799853, rs1057910	(58)	10	(+)	96,588,430 96,839,137	rs9332235	96,737,895	CYP2C9	G/A	0.025	-0.011	0.0043	204 (44)	0.1887	0.5943
DAB2IP	rs7025486	(59)	9	(+)	123,269,220 123,687,628	rs13290547	123,527,316	DAB2IP	C/T	0.063	-0.009	3.6E-05	363 (97)	0.0035	0.0665
ENPP1	K121Q	(60)	6	(+)	132,070,853 132,354,043	rs11154647	132,254,445	ENPP1	G/T	0.052	-0.007	0.0034	324 (86)	0.2897	0.7598
SELE (E-selectin)	S128R	(53,61)	1	(-)	167,858,406 168,069,803	rs7531806	167,917,668	SELL	G/A	0.561	-0.003	0.0045	348 (65)	0.2901	0.7598
F13A1	V34L	(62)	6	(-)	5,989,310 6,365,923	rs11756027	6,077,751	F13A1	C/T	0.756	0.003	0.0109	638 (211)	1.0000	1.0000
F2	Prothrombin G20210A	(63-66)	11	(+)	46,597,331 46,817,631	rs7109698	46,629,920	KIAA0652	C/T	0.170	-0.004	0.0009	83 (17)	0.0155	0.2132
F5	G1691A (Leiden)	(63-66)	1	(-)	167,647,816 167,922,393	rs2040444	167,749,060	F5	G/A	0.525	0.003	0.0024	441 (105)	0.2534	0.7334
F7	AA I/D at 11293, R353q	(67-69)	13	(+)	112,708,106 112,922,995	rs2993312	112,731,466	MCF2L	G/A	0.645	-0.003	0.0106	179 (71)	0.7491	1.0000
FGB (FBG)	BcI I digestion; T1689G; G-455A	(67,69-72)	4	(+)	155,603,596 155,811,686	rs6054	155,709,058	FGB	C/T	0.006	-0.048	0.0007	174 (57)	0.0384	0.2740
FGG	10034C>T, rs2066865	(73)	4	(-)	155,644,737 155,853,352	rs6054	155,709,058	FGB	C/T	0.006	-0.048	0.0007	163 (55)	0.0371	0.2740
F12 (FXII)	rs17876008 (4C>T)	(74)	5	(-)	176,661,747 176,869,183	rs335467	176,823,248	DBN1	G/A	0.457	0.003	0.0455	65 (24)	1.0000	1.0000
GNB3	825C>T	(46)	12	(+)	6,719,636 6,926,817	rs729751	6,924,611	C12orf57	C/T	0.039	-0.010	0.0034	137 (57)	0.1945	0.5943
GSTM1		(75)	1	(+)	109,931,965 110,137,889	rs655315	110,016,701	GSTM2	G/A	0.497	0.005	0.0004	163 (63)	0.0277	0.2740
GSTT1		(75)	22	(-)	22,606,141 22,814,231	rs9624387	22,750,128	CABIN1	C/T	0.075	-0.004	0.0269	64 (24)	0.6444	1.0000

to be continued

Supplementary Table 10: continued

						Most strongly associated SNP within the gene region ± 100 kb upstream and downstream of the candidate gene								
Candidate gene	Investigated markers as reported	Ref.	Chr.	Coding strand	Gene region 5' start 3' end	Top hit rs number	Position (bp)	Nearest gene	Ref./ coded allele	Coded allele freq.	Effect	P-value	# of loci (eff. loci)*	corr. P-value**	False discovery rate‡
HFE	C282Y, H63D	(76)	6	(+)	26,095,488 26,305,035	rs11751812	26,295,313	HIST1H4D	G/T	0.095	0.004	0.0177	183 (48)	0.8472	1.0000
HIF1A	P582S, rs11549465 A588T, rs11549467	(77)	14	(+)	61,131,992 61,384,729	rs17098997	61,171,460	HIF1A	G/C	0.192	0.004	0.0010	206 (51)	0.0509	0.3113
HP (hapto-globin)	Hp1, Hp2	(78)	16	(+)	70,546,009 70,752,458	rs16973520	70,553,792	KIAA0174	G/T	0.748	0.004	0.0007	204 (54)	0.0399	0.2740
ICAM1	K469E	(52,53, 61)	19	(+)	10,142,779 10,358,291	rs281419	10,277,842	GLP-1	G/A	0.149	-0.003	0.1610	113 (43)	1.0000	1.0000
IL1B	+3953	(79)	2	(-)	113,203,808 113,410,827	rs12472089	113,371,320	IL1F7	C/T	0.326	0.003	0.0168	192 (34)	0.5705	1.0000
IL6	-174 G/C	(52,53, 80)	7	(+)	22,633,345 22,838,141	rs7796691	22,768,932	IL6	C/T	0.797	-0.003	0.0089	261 (73)	0.6516	1.0000
ITGB3	PLA1/A2; $P1^A$	(71,81)	17	(+)	42,586,207 42,845,075	rs11871477	42,783,505	C17orf57	G/A	0.985	0.012	0.0173	206 (53)	0.9164	1.0000
LIPC	rs2070895 (G-250A)	(82)	15	(+)	56,411,467 56,748,364	rs261292	56,468,196	LIPC	G/A	0.004	0.047	0.0027	500 (178)	0.4758	1.0000
LPA	K-IV repeats; rs1853021 (C93T)	(83,84)	6	(-)	160,772,506 161,107,397	rs10455872	160,930,108	LPA	G/A	0.936	0.007	0.0062	268 (74)	0.4615	1.0000
CCL2 (MCP1)	-2518 A/G	(53)	17	(+)	29,506,409 29,708,331	rs2368697	29,522,313	ACCN1	G/A	0.257	-0.002	0.0447	236 (64)	1.0000	1.0000
MIF	-173 G/C	(53)	22	(+)	22,466,565 22,667,409	rs17004046	22,570,304	MIF	G/T	0.145	-0.003	0.0580	211 (57)	1.0000	1.0000
MMP1	-1607 1G/2G	(53)	11	(-)	102,065,861 102,274,104	rs11225443	102,248,161	MMP12	C/T	0.033	-0.011	0.0081	329 (90)	0.7254	1.0000
MMP3	-1171 5A/6A	(53)	11	(-)	102,111,738 102,319,552	rs11225443	102,248,161	MMP12	C/T	0.033	-0.011	0.0081	257 (79)	0.6367	1.0000
MMP9	-1563 C/T	(53)	20	(+)	43,970,954 44,178,606	rs11698938	44,171,037	CD40	C/T	0.010	-0.026	0.0026	174 (53)	0.1389	0.5877

to be continued

Supplementary Table 10: continued

						Most strongly associated SNP within the gene region ± 100 kb upstream and downstream of the candidate gene									
Candidate gene	Investigated markers as reported	Ref.	Chr.	Coding strand	Gene region 5' start 3' end	Top hit rs number	Position (bp)	Nearest gene	Ref./ coded allele	Coded allele freq.	Effect	P-value	# of loci (eff. loci)*	corr. P-value**	False discovery rate‡
MTHFR	C677T	(65,66, 85)	1	(-)	11,668,374 11,888,702	rs2050265	11,802,286	CLCN6	G/A	0.840	0.003	0.0286	221 (62)	1.0000	1.0000
MTTP (MTP)	-493G>T	(86)	4	(+)	100,615,004 100,863,649	rs11730739	100,859,020	MTTP	C/T	0.291	-0.003	0.0163	216 (46)	0.7480	1.0000
NOS3	14 SNPs; -786T>C, -894G>T, 4a/4b	(43,87)	7	(+)	150,219,080 150,442,608	rs11763819	150,247,598	KCNH2	T/A	0.975	-0.010	0.0175	127 (47)	0.8239	1.0000
P2RY12	H2 haplotype	(88)	3	(-)	152,438,066 152,685,234	rs3821665	152,616,506	MED12L	G/A	0.514	0.003	0.0391	294 (60)	1.0000	1.0000
PLA2G7 (PAF-AH)	G994T in exon 9	(89)	6	(-)	46,680,238 46,911,055	rs1833460	46,885,336	MEP1A	C/T	0.112	0.006	0.0037	205 (47)	0.1719	0.5943
SERPINE1 (PAI1)	HindIII	(67)	7	(+)	100,457,172 100,669,026	rs4727479	100,552,387	SERPINE1	G/C	0.130	-0.005	0.0039	131 (47)	0.1851	0.5943
PPARG	rs1801282 (Pro12Ala)	(90)	3	(+)	12,204,349 12,550,854	rs310766	12,208,482	SYN2	G/A	0.254	-0.002	0.0515	314 (82)	1.0000	1.0000
SCARB1	rs4238001, intron 5 (C>T at intron pos.55), rs5888	(91)	12	(-)	123,728,129 124,014,287	rs11057830	123,873,006	SCARB1	G/A	0.154	-0.004	0.0069	265 (99)	0.6828	1.0000
SLC2A10	10 SNPs	(92)	20	(+)	44,671,686 44,898,390	rs11550540	44,749,176	TP53RK	C/T	0.898	0.005	0.0081	230 (77)	0.6208	1.0000
SRD5A1	HinfI SNP	(93)	5	(+)	6,586,500 6,822,673	rs12522035	6,819,185	POLS	C/A	0.926	-0.004	0.0318	346 (101)	1.0000	1.0000
SRD5A2	(V89L	(93)	2	(-)	31,503,160 31,759,544	rs7424544	31,625,875	SRD5A2	C/T	0.977	-0.007	0.0441	127 (25)	1.0000	1.0000
UGT1A1	-53 TA-repeat	(33)	2	(+)	234,233,658 234,446,684	rs11690786	234,357,356	UGT1A1	C/T	0.315	0.003	0.0036	597 (137)	0.4924	1.0000
VKORC1		(94)	16	(-)	30,909,677 31,113,777	rs17708638	31,078,675	PRSS36	C/T	0.310	0.003	0.0061	54 (12)	0.0736	0.4048
ZNF202	rs10893081 (-660A>G)	(95)	11	(-)	123,000,207 123,217,573	rs1148107	123,020,997	SCN3B	C/A	0.969	0.011	0.0015	308 (70)	0.1064	0.4877

to be continued

PubMed search terms "((ankle-brachial index) OR (peripheral arterial disease)) AND polymorphism"

For all candidate genes the official HGNC-approved gene names are reported and, where needed, the alias names used in the publications are provided in brackets. The chromosomal coordinates refer to the plus-strand according to HapMap Rel 24 (based on NCBI Genome Build 36 and dbSNP b.126). In the case that more than one RefSeq transcript is known, the +/- 100 kb gene region was defined starting from the coordinates which comprised all known transcripts (i.e. starting from the beginning of the outmost 5' exon to the end of the outmost 3' exon). The strand actually encoding the candidate gene is reported in the column "coding strand".

* Effective number of independent loci calculated using the function reported by Gao [(38)] based on imputed genotype data in KORA F4
** P-value Bonferroni corrected within each gene region using the effective number of loci
*** Chromosome 11 region: Chromosomal region identified by admixture mapping
‡ Using the corrected p-values of the SNP in association to ABI, we calculated a false discovery rate (FDR) to account for the number of gene regions examined. An FDR <0.10 defined evidence of a significant association.

Supplementary Table 11: Candidate SNPs for myocardial infarction and /or coronary artery disease and their association with ABI in the CHARGE GWAS discovery sample (population isolates excluded) †.

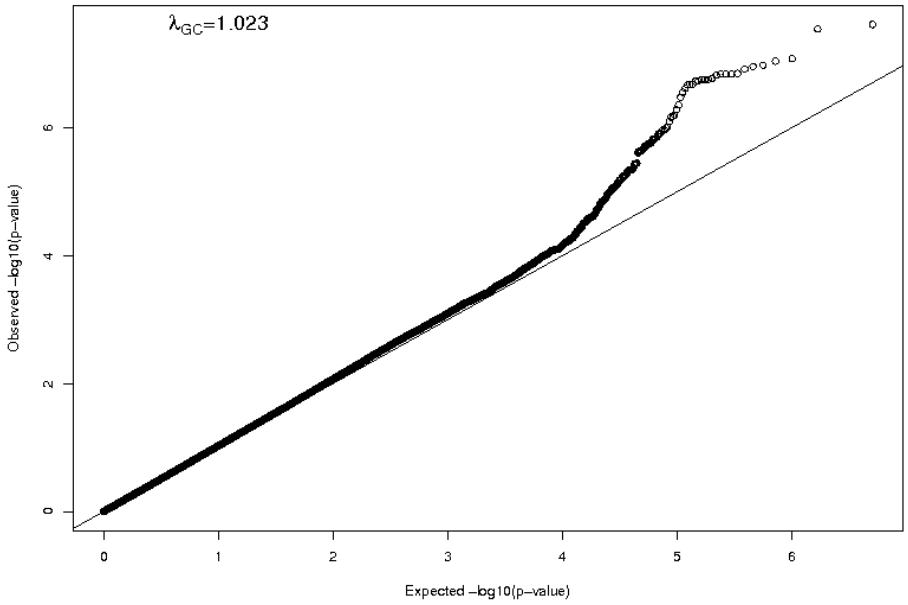
SNP	Chr	Physical Position	Closest Gene	Risk / Non-risk Allele	Risk Allele frequency	N	Beta	SE	P value	False discovery rate‡
rs4977574	9	22,088,574	*CDKN2A, CKDN2B*	G/A	0.49	35411	-0.0047	0.001	2.33E-06	6.52E-05
rs1122608	19	11,024,601	*LDLR*	G/T	0.74	35384	-0.0035	0.001	2.56E-03	0.036
rs1412444	10	90,992,907	*LIPA*	C/T	0.67	26195	0.0027	0.001	3.05E-02	0.284
rs2505083	10	30,375,128	*KIAA1462*	C/T	0.42	35411	0.0018	0.001	8.28E-02	0.58
rs46522	17	44,343,596	*UBE2Z,GIP,ATP5G1,SNF8*	T/C	0.54	35364	-0.0015	0.001	1.17E-01	0.594
rs12936587	17	17,484,447	*RASD1, SMCR3, PEMT*	G/A	0.55	35346	0.0016	0.001	1.27E-01	0.594
rs2895811	14	99,203,695	*HHIPL1*	C/T	0.42	35396	0.0014	0.001	1.76E-01	0.69
rs964184	11	116,154,127	*ZNF259,APOA5-A4-C3-A1*	G/C	0.14	35384	0.0019	0.002	1.97E-01	0.69
rs3184504	12	110,368,991	*SH2B3*	T/C	0.5	35359	-0.0011	0.001	2.70E-01	0.762
rs2306374	3	139,602,642	*MRAS*	C/T	0.16	35410	0.0015	0.001	2.81E-01	0.762
rs6903956	6	11,882,569	*C6orf105*	G/A	0.63	30183	0.0012	0.001	2.99E-01	0.762
rs12413409	10	104,709,086	*CYP17A1,CNNM2,NT5C2*	G/A	0.91	35413	0.0016	0.002	3.61E-01	0.843
rs10953541	7	107,031,781	*7q22*	C/T	0.75	35369	0.0009	0.001	4.52E-01	0.886
rs974819	11	103,165,777	*PDGFD*	C/T	0.71	35280	0.0007	0.001	5.05E-01	0.886
rs4773144	13	109,758,713	*COL4A1,COL4A2*	A/G	0.55	33491	-0.0007	0.001	5.06E-01	0.886
rs12190287	6	134,256,218	*TCF21*	G/C	0.38	35168	0.0007	0.001	5.32E-01	0.886
rs6725887	2	203,454,130	*WDR12*	C/T	0.13	35399	0.0009	0.002	5.54E-01	0.886
rs17114036	1	56,735,409	*PPAP2B*	A/G	0.9	35187	-0.001	0.002	5.91E-01	0.886
rs9982601	21	34,520,998	*MRPS6*	T/C	0.14	35343	-0.0008	0.002	6.01E-01	0.886
rs17609940	6	35,142,778	*ANKS1A*	G/C	0.8	35382	0.0006	0.001	6.51E-01	0.911
rs216172	17	2,073,254	*SMG6, SRR*	G/C	0.65	26051	0.0004	0.001	7.42E-01	0.927
rs1746048	10	44,095,830	*CXCL12*	T/C	0.13	35403	-0.0004	0.002	7.59E-01	0.927
rs3825807	15	76,876,166	*ADAMTS7*	G/A	0.44	35338	0.0003	0.001	7.61E-01	0.927
rs579459	9	135,143,989	*ABO*	C/T	0.22	35371	0.0002	0.001	8.72E-01	0.98
rs12526453	6	13,035,530	*PHACTR1*	C/G	0.65	35353	-0.0001	0.001	9.09E-01	0.98
rs11206510	1	55,268,627	*PCSK9*	T/C	0.81	35174	-0.0001	0.001	9.54E-01	0.98
rs599839	1	109,623,689	*SORT1*	G/A	0.23	35379	0	0.001	9.75E-01	0.98
rs11556924	7	129,450,732	*ZC3HC1*	C/T	0.61	35217	0	0.001	9.80E-01	0.98

† Candidate genes for look-up of association with ABI were identified by recent GWAS to be genome-wide significantly associated with CAD.

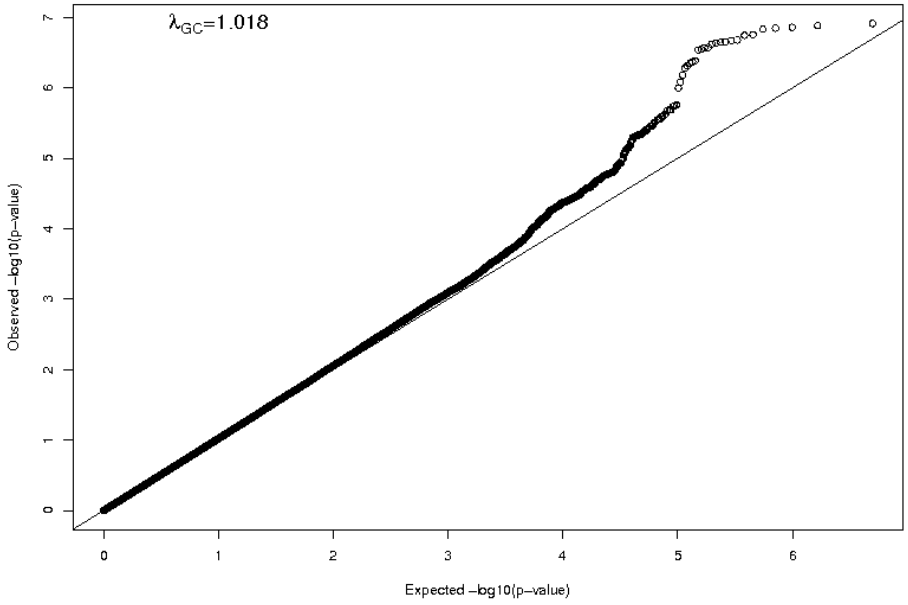
‡ Using the p-values of the SNP in association to ABI, we calculated a false discovery rate (FDR) to account for the number of gene regions examined. An FDR <0.10 defined evidence of a significant association.

Suppl Figure 1: **Quantile-quantile plots of the meta-analysis of GWAS of ABI.** The dots represent the observed -log10 p-values associated with ABI. The expected distribution of -log10 p-values under the null hypothesis is shown by the straight line. A) excluding the population isolates; B) including the population isolates

A

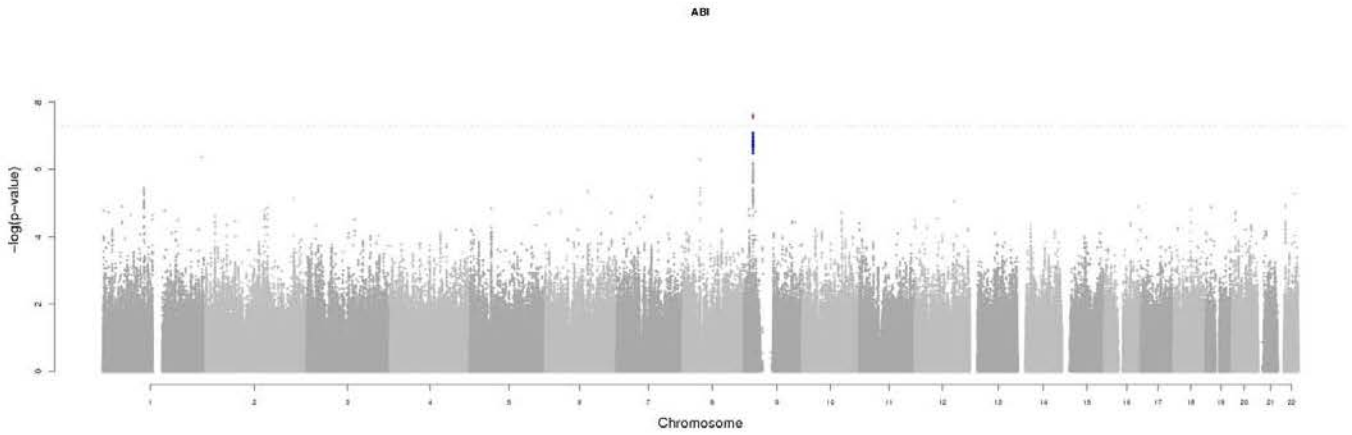


B

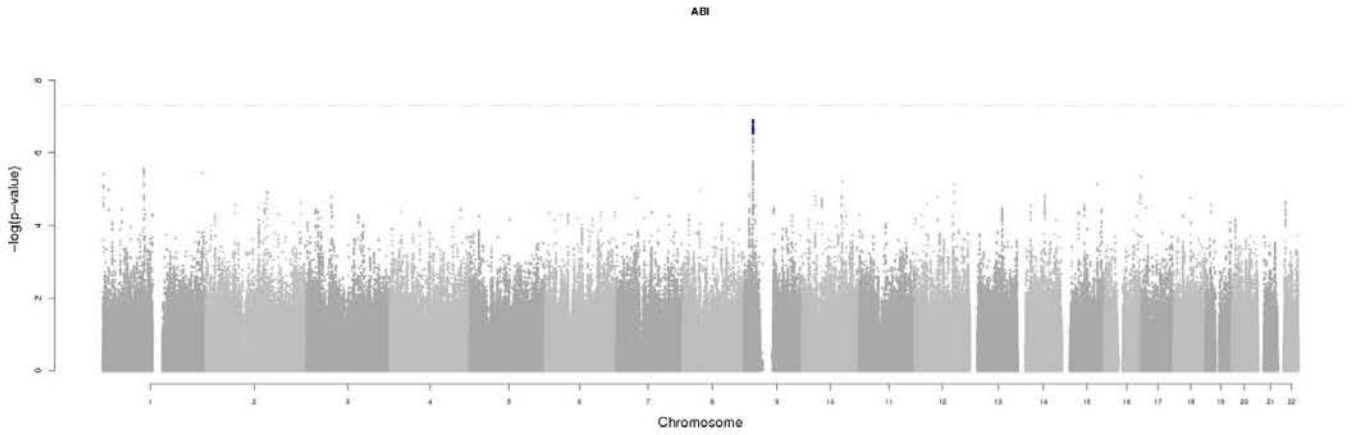


Suppl Figure 2: Manhattan plots of the meta-analysis of GWAS of ABI. X-axis represents the chromosomal position for each SNP, and the y-axis the -log10 p-value for association with ABI. The horizontal dotted line represents the genome-wide significance level of P<5x10-8. A) excluding the population isolates; B) including the population isolates

A



B



Acknowledgments/Funding Sources

Study	Acknowledgements/Funding
Amish Study (Amish)	This work was supported by research grants R01 088119, R01 AR046838, U01 HL72515, and R01 AG18728. Additional support was provided by the University of Maryland General Clinical Research Center, Grant M01 RR 16500, Mid-Atlantic Nutrition Obesity Research Center Grant P30 DK072488, General Clinical Research Centers Program, National Center for Research Resources (NCRR), NIH, and the Baltimore Veterans Administration Geriatric Research and Education Clinical Center (GRECC). Dr. Montasser was supported by T32 training Grant AG000219. We gratefully acknowledge our Amish liaisons and field workers and the extraordinary cooperation and support of the Amish community without which these studies would not have been possible.
Atherosclerosis Risk in Communities Study (ARIC)	The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.
Cardiovascular Health Study (CHS)	This CHS research was supported by NHLBI contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133 and NHLBI grants HL080295, HL075366, HL087652, HL105756 with additional contribution from NINDS. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIA. See also http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01-RR00425 to the Cedars-Sinai General Clinical Research Center Genotyping core, Cedars-Sinai Board of Governors' Chair in Medical Genetics (JIR), and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.
Family Heart Study (FamHS)	The Family Heart Study GWAS was funded by gran HL08770002 and the work was supported by contract numbers R01HL08770003 from the NHLBI, and R01DK06833603 and R01DK07568101 from NIDDK.
Framingham Heart Study (FHS)	This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of

	this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.
Genetic Epidemiology Network of Arteriopathy (GENOA)	"Genetic Epidemiology Network of Arteriopathy (GENOA) study is supported by the National Institutes of Health, grant number 5R01HL087660 from National Heart, Lung and Blood Institute."
Gutenberg Heart Study (GHS)	The Gutenberg Heart Study is funded through the government of Rheinland-Pfalz ("Stiftung Rheinland Pfalz für Innovation", contract number AZ 961-386261/733), the research programs "Wissen schafft Zukunft" and "Schwerpunkt Vaskuläre Prävention" of the Johannes Gutenberg-University of Mainz and its contract with Boehringer Ingelheim and PHILIPS Medical Systems including an unrestricted grant for the Gutenberg Heart Study. This research was also supported by the National Genome Network "NGFNplus" (contract number project A3 01GS0833 and 01GS0831) by the Federal Ministry of Education and Research, Germany
Health, Aging and Body Composition (Health ABC)	This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C.
Invecchiare in Chianti Study (InCHIANTI)	The InCHIANTI study baseline (1998-2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336); the InCHIANTI Follow-up 1 (2001-2003) was funded by the U.S. National Institute on Aging (Contracts: N.1-AG-1-1 and N.1-AG-1-2111); the InCHIANTI Follow-ups 2 and 3 studies (2004-2010) were financed by the U.S. National Institute on Aging (Contract: N01-AG-5-0002);supported in part by the Intramural research program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland. This research was supported in part by the Intramural Research Program of the National Institute on Aging. A portion of the support was through a R&D contract with MedStar Health Research Institute.
Cooperative Health Research in the Region of Augsburg (KORA F3, KORA F4)	This study was partially funded by the "Genomics of Lipid-associated Disorders - GOLD" of the "Austrian Genome Research Programme GEN-AU" and by the Austrian Heart Fund to F. Kronenberg and by the Austrian National Bank (Project-Nr. 13662) to Barbara Kollerits. The MONICA/KORA Augsburg cohort study was financed by the Helmholtz Zentrum München and the German National Genome Research Net NGFN2 and NGFNplus (H.-E. Wichmann 01GS0823).
Netherlands Study of Anxiety and Depression (NESDA)	NESDA was supported by the Geestkracht program of ZonMW [grant 10-000-1002]; matching funds from universities and mental health care institutes involved in NESDA (GGZ Buitenamstel-Geestgronden, Rivierduinen, University Medical Center Groningen, GGZ 25 Lentis, GGZ Friesland, GGZ Drenthe). Genotyping was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health, and analysis was supported by grants from GAIN and

	the NIMH (MH081802). Genotype data were obtained from dbGaP (http://www.ncbi.nlm.nih.gov/dbgap, accession number phs000020.v1.p1). Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.
Nijmegen Biomedical Study (NBS)	The Nijmegen Biomedical Study (NBS) was initiated in 2000 as a large study among the inhabitants of the municipality of Nijmegen by the departments of Epidemiology, Biostatistics and HTA, Clinical Chemistry, and Endocrinology of the Radboud University Nijmegen Medical Centre (RUNMC) in collaboration with the municipality of Nijmegen en de community health service of Nijmegen. Support was obtained from RUNMC. The measurements of ABI and other non-invasive measurements of atherosclerosis (NIMA) was supported by Grant 2003B057 of the Netherlands Heart Foundation. This work was sponsored by the Stichting Nationale Computerfaciliteiten (National Computing Facilities Foundation, NCF) for the use of supercomputer facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research, NWO).
Rotterdam Study (RS-I and RS-II)	The Rotterdam GWA study was funded by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Consortium for Healthy Aging (NCHA) project nr. 050-060-810. This study is further supported by NWO grant (vici, 918-76-619). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are very grateful to the participants and staff from the Rotterdam Study, the participating general practioners and the pharmacists. We thank Pascal Arp, Mila Jhamai, Dr Michael Moorhouse, MarijnVerkerk, and Sander Bervoets for their help in creating the GWAS database. We would like to thank Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German BundesministeriumfuerForschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 G, for access to their grid resources.
Study of Health in Pomerania (SHIP)	SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG.
CROATIA Study (Croatia-Vis,	The CROATIA study would like to acknowledge the invaluable contributions of the recruitment teams (including those from the Institute of Anthropological Research

Croatia-Korcula, Croatia-Split)	in Zagreb) in Vis and Korcula and at the Croatian Centre for Global Health at the University of Split, the administrative teams in Croatia and Edinburgh and the people of Vis, Korcula and Split. Genotyping was performed at the Wellcome Trust Clinical Research Facility in Edinburgh for CROATIA-Vis, by Helmholtz Zentrum München, GmbH, Neuherberg, Germany for CROATIA-Korcula and by AROS Applied Biotechnology, Aarhus, Denmark for CROATIA-Split. These studies were supported by grants from the Medical Research Council UK and Ministry of Science, Education and Sport of the Republic of Croatia (No. 108-1080315-0302) and CROATIA-Vis by the European Union framework program 6 European Special Populations Research Network (EUROSPAN) project (contract LSHG-CT-2006-018947)
Erasmus Rucphen Family Study (ERF)	The genotyping for the ERF study was supported by EUROSPAN (European Special Populations Research Network) and the European Commission FP6 STRP grant (018947; LSHG-CT-2006-01947). The ERF study was further supported by grants from the Netherlands Organisation for Scientific Research, Erasmus MC, the Centre for Medical Systems Biology (CMSB) and the Netherlands Brain Foundation (HersenStichting Nederland). We are grateful to all patients and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection.
Orkney Complex Disease Study (Orcades)	ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.
Bruneck Study (Bruneck)	The study was supported by the Pustertaler Verein zur Prävention von Herz- und Hirngefaesserkrankungen, Gesundheitsbezirk Bruneck, and the Assessorat fuer Gesundheit, Province of Bolzano, Italy.
Copenhagen City Heart Study (CCHS)	This work was supported by a Specific Targeted Research Project grant from the European Union, Sixth Framework Programme Priority (FP-2005-LIFESCIHEALTH-6) contract 037631, the Danish Medical Research Council (Copenhagen), the Research Fund at Rigshospitalet, Copenhagen University Hospital (Copenhagen), Chief Physician Johan Boserup and Lise Boserup's Fund (Copenhagen), Ingeborg and Leo Dannin's Grant (Copenhagen), and Henry Hansen's and Wife's Grant (Copenhagen). Genotyping was supported by a grant from the Austrian Heart Fund to F. Kronenberg.
National Health and Nutrition Examination Surveys (NHANES)	The study participants derive from the National Health and Nutrition Examination Surveys (NHANES), and these studies are supported by the Centers for Disease Control and Prevention. We would like to thank Dr. Geraldine McQuillan and Jody McLean for their help in accessing the Genetic NHANES data. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention. The Vanderbilt University Center for Human Genetics Research, Computational Genomics Core provided computational and/or analytical support for this work. Genotyping was performed in the Vanderbilt DNA Resources Core by Paxton Baker under the direction of Dr. Holli Dilks and Cara Sutcliffe.
Multi-Ethnic Study of	MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators.

Atherosclerosis (MESA)	Support is provided by grants and contracts N01 HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and RR-024156. MESA SNP Health Association Resource (SHARe): “Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278.”
Prevention of Renal and Vascular End-stage Disease (PREVEND)	The PREVEND study is supported by the Dutch Kidney Foundation (Grant E033), EU project grant GENECURE (FP-6 LSHM CT 2006 037697), and NWO VENI (grant number 916.76.170)
Cardiovascular Disease in Intermittent Claudication (CAVASIC)	This study was partially funded by the "Genomics of Lipid-associated Disorders - GOLD" of the "Austrian Genome Research Programme GEN-AU" and by the Austrian Heart Fund to F. Kronenberg and by the Austrian National Bank (Project-Nr. 13662) to Barbara Kollerits.
Genetic Determinants of Peripheral Arterial Disease (GenePAD)	This study was supported by grants from the National Heart, Lung and Blood Institute (RO1 HL-75774, 1K12 HL087746, 1P50HL083800) as well as a National Institutes of Health grant M01 RR 00070 (General Clinical Research Center, Stanford University School of Medicine) and the Stanford Cardiovascular Institute.
Linz Peripheral Arterial Disease (LIPAD)	This project was supported in part by a grant for scientific research from the Upper Austrian Government. Genotyping was supported by a grant from the Austrian Heart Fund to F. Kronenberg.

References

1. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am. J. Epidemiol. 129:687-702, 1989.
2. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A, .: The Cardiovascular Health Study: design and rationale. Ann. Epidemiol. 1:263-276, 1991.
3. Higgins M, Province M, Heiss G, Eckfeldt J, Ellison RC, Folsom AR, Rao DC, Sprafka M, Williams R: NHLBI Family Heart Study: objectives and design. Am. J. Epidemiol. 143:1219-1228, 1996.
4. Dawber TR, MEADORS GF, MOORE FE, Jr.: Epidemiological approaches to heart disease: the Framingham Study. Am. J. Public Health Nations. Health 41:279-281, 1951.
5. Dawber TR, Kannel WB, LYELL LP: An approach to longitudinal studies in a community: the Framingham Study. Ann. N. Y. Acad. Sci. 107:539-556, 1963.
6. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP: The Framingham Offspring Study. Design and preliminary data. Prev. Med. 4:518-525, 1975.
7. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP: An investigation of coronary heart disease in families. The Framingham offspring study. Am. J. Epidemiol. 110:281-290, 1979.
8. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, Jr., Jones DW, Materson BJ, Oparil S, Wright JT, Jr., Roccella EJ: The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA 289:2560-2572, 2003.
9. Ferrucci L, Bandinelli S, Benvenuti E, Di IA, Macchi C, Harris TB, Guralnik JM: Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. J. Am. Geriatr. Soc. 48:1618-1625, 2000.
10. Wichmann HE, Gieger C, Illig T, for the MONIKA/KORA Study Group: KORA-gen: Resource for Population Genetics, Controls and a Broad Spectrum of Disease Phenotypes. Gesundheitswesen 67:S26-S30, 2005.
11. Penninx BW, Beekman AT, Smit JH, et al: The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. Int. J. Methods Psychiatr. Res. 17:121-140, 2008.
12. Sullivan PF, De Geus EJ, Willemsen G, et al: Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol. Psychiatry 14:359-375, 2009.
13. Hoogendoorn EH, Hermus AR, de VF, Ross HA, Verbeek AL, Kiemeney LA, Swinkels DW, Sweep FC, den HM: Thyroid function and prevalence of anti-thyroperoxidase antibodies in a population with borderline sufficient iodine intake: influences of age and sex. Clin. Chem. 52:104-111, 2006.
14. Kiemeney LA, Thorlacius S, Sulem P, et al: Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat. Genet 40:1307-1312, 2008.
15. Holewijn S, den HM, Swinkels DW, Stalenhoef AF, De GJ: The metabolic syndrome and its traits as risk factors for subclinical atherosclerosis. J. Clin. Endocrinol. Metab 94:2893-2899, 2009.
16. Hofman A, Breteler MM, Van Duijn CM, Janssen HL, Krestin GP, Kuipers EJ, Stricker BH, Tiemeier H, Uitterlinden AG, Vingerling JR, Witteman JC: The Rotterdam Study: 2010 objectives and design update. Eur. J. Epidemiol. 24:553-572, 2009.
17. John U, Greiner B, Hensel E, et al: Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. Soz. Praventivmed. 46:186-194, 2001.
18. Volzke H, Alte D, Schmidt CO, et al: Cohort profile: the study of health in Pomerania. Int. J. Epidemiol. 40:294-307, 2011.
19. Mitchell BD, McArdle PF, Shen H, et al: The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. Am. Heart J. 155:823-828, 2008.

20. Post W, Bielak LF, Ryan KA, Cheng YC, Shen H, Rumberger JA, Sheedy PF, Shuldiner AR, Peyser PA, Mitchell BD: Determinants of coronary artery and aortic calcification in the Old Order Amish. Circulation 115:717-724, 2007.
21. Aulchenko YS, Heutink P, Mackay I, Bertoli-Avella AM, Pullen J, Vaessen N, Rademaker TA, Sandkuijl LA, Cardon L, Oostra B, Van Duijn CM: Linkage disequilibrium in young genetically isolated Dutch population. Eur. J. Hum. Genet 12:527-534, 2004.
22. Kiechl S, Willeit J, for the Bruneck Study Group: The natural course of atherosclerosis. Part I: incidence and progression. Arterioscler. Thromb. Vasc. Biol. 19:1484-1490, 1999.
23. Kiechl S, Willeit J, for the Collaborative Study Group: The natural course of atherosclerosis. Part II: vascular remodeling. Arterioscler. Thromb. Vasc. Biol. 19:1491-1498, 1999.
24. Kiechl S, Lorenz E, Reindl M, Wiedermann CJ, Oberhollenzer F, Bonora E, Willeit J, Schwartz DA: Toll-like receptor 4 polymorphisms and atherogenesis. N. Engl. J. Med. 347:185-192, 2002.
25. Kronenberg F, Kronenberg MF, Kiechl S, Trenkwalder E, Santer P, Oberhollenzer F, Egger G, Utermann G, Willeit J: Role of lipoprotein(a) and apolipoprotein(a) phenotype in atherogenesis: prospective results from the Bruneck Study. Circulation 100:1154-1160, 1999.
26. Willeit J, Kiechl S, Oberhollenzer F, Rungger G, Egger G, Bonora E, Mitterer M, Muggeo M: Distinct risk profiles of early and advanced atherosclerosis : prospective results from the bruneck study. Arterioscler. Thromb. Vasc. Biol. 20:529-537, 2000.
27. Benn M, Nordestgaard BG, Jensen GB, Tybjaerg-Hansen A: Improving prediction of ischemic cardiovascular disease in the general population using apolipoprotein B: the Copenhagen City Heart Study. Arterioscler Thromb Vasc Biol 27:661-670, 2007.
28. Frikke-Schmidt R, Sing CF, Nordestgaard BG, Steffensen R, Tybjaerg-Hansen A: Subsets of SNPs define rare genotype classes that predict ischemic heart disease. Hum Genet 120:865-877, 2007.
29. Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, Hobbs HH, Cohen JC: Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat Genet 39:513-516, 2007.
30. Bild DE, Bluemke DA, Burke GL, et al: Multi-ethnic study of atherosclerosis: objectives and design. Am. J. Epidemiol. 156:871-881, 2002.
31. McQuillan GM, Porter KS, Agelli M, Kington R: Consent for genetic research in a general population: the NHANES experience. Genet Med 5:35-42, 2003.
32. McQuillan GM, Pan Q, Porter KS: Consent for genetic research in a general population: an update on the National Health and Nutrition Examination Survey experience. Genet Med 8:354-360, 2006.
33. Rantner B, Kollerits B, Anderwald-Stadler M, Klein-Weigel P, Gruber I, Gehringer A, Haak M, Schnapka-Köpf M, Fraedrich G, Kronenberg F: Association between the *UGT1A1* TA-repeat polymorphism and bilirubin concentration in patients with intermittent claudication: results from the CAVASIC Study. Clin. Chem. 54:851-857, 2008.
34. Sacks D, Bakal CW, Beatty PT, Becker GJ, Cardella JF, Raabe RD, Wiener HM, Lewis CA: Position statement on the use of the ankle brachial index in the evaluation of patients with peripheral vascular disease. A consensus statement developed by the Standards Division of the Society of Interventional Radiology. J. Vasc. Interv. Radiol. 14:S389, 2003.
35. Chen YH, Liu CK, Chang SC, Lin YJ, Tsai MF, Chen YT, Yao A: GenoWatch: a disease gene mining browser for association study. Nucleic Acids Res. 36:W336-W340, 2008.
36. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. U. S. A 106:9362-9367, 2009.
37. Yu W, Clyne M, Khoury MJ, Gwinn M: Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations. Bioinformatics 26:145-146, 2010.
38. Gao X: Multiple testing corrections for imputed SNPs. Genet Epidemiol. 35:154-158, 2011.

39. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. , Series B 57:289-300, 1995.
40. Renner W, Pabst E, Paulweber B, Malaimare L, Iglseder B, Wascher TC, Pilger E: The angiotensin-converting-enzyme insertion/deletion polymorphism is not a risk factor for peripheral arterial disease. Atherosclerosis 165:175-178, 2002.
41. Schelleman H, Klungel OH, Witteman JC, Hofman A, Van Duijn CM, de BA, Stricker BH: Pharmacogenetic interactions of three candidate gene polymorphisms with ACE-inhibitors or beta-blockers and the risk of atherosclerosis. Br. J. Clin. Pharmacol. 64:57-66, 2007.
42. Li R, Nicklas B, Pahor M, Newman A, Sutton-Tyrrell K, Harris T, Lakatta E, Bauer DC, Ding J, Satterfield S, Kritchevsky SB: Polymorphisms of angiotensinogen and angiotensin-converting enzyme associated with lower extremity arterial disease in the Health, Aging and Body Composition study. J. Hum. Hypertens. 21:673-682, 2007.
43. Sticchi E, Sofi F, Romagnuolo I, Pratesi G, Pulli R, Pratesi C, Abbate R, Fatini C: eNOS and ACE genes influence peripheral arterial disease predisposition in smokers. J. Vasc. Surg. 52:97-102, 2010.
44. Fatini C, Sticchi E, Sofi F, Said AA, Pratesi G, Pulli R, Pratesi C, Abbate R: Multilocus analysis in candidate genes ACE, AGT, and AGTR1 and predisposition to peripheral arterial disease: role of ACE D/-240T haplotype. J. Vasc. Surg. 50:1399-1404, 2009.
45. Park EY, Ahn HM, Lee JA, Hong YM: Insertion/deletion polymorphism of angiotensin converting enzyme gene in Korean hypertensive adolescents. Heart Vessels 24:193-198, 2009.
46. Morrison AC, Bray MS, Folsom AR, Boerwinkle E: ADD1 460W allele associated with cardiovascular disease in hypertensive individuals. Hypertension 39:1053-1057, 2002.
47. Hindorff LA, Heckbert SR, Psaty BM, Lumley T, Siscovick DS, Herrington DM, Edwards KL, Tracy RP: beta(2)-Adrenergic receptor polymorphisms and determinants of cardiovascular risk: the Cardiovascular Health Study. Am. J. Hypertens. 18:392-397, 2005.
48. Tsai PC, Lin TH, Hsu PC, Wang YS, Liao YC, Juo SH: Polymorphism of 270 A>G in BRAP is Associated with Lower Ankle-Brachial Index in a Taiwanese Population. J. Atheroscler. Thromb. 0:2011.
49. Cluett C, McDermott MM, Guralnik J, et al: The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people. Circ. Cardiovasc. Genet 2:347-353, 2009.
50. Scherer ML, Nalls MA, Pawlikowska L, et al: Admixture mapping of ankle-arm index: identification of a candidate locus associated with peripheral arterial disease. J. Med. Genet 47:1-7, 2010.
51. Thorgeirsson TE, Geller F, Sulem P, et al: A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 452:638-642, 2008.
52. Zintzaras E, Zdoukopoulos N: A field synopsis and meta-analysis of genetic association studies in peripheral arterial disease: The CUMAGAS-PAD database. Am. J. Epidemiol. 170:1-11, 2009.
53. Flex A, Gaetani E, Angelini F, Sabusco A, Chilla C, Straface G, Biscetti F, Pola P, Castellot JJ, Jr., Pola R: Pro-inflammatory genetic profiles in subjects with peripheral arterial occlusive disease and critical limb ischemia. J. Intern. Med. 262:124-130, 2007.
54. Pitha J, Hubacek JA, Pithova P: The connexin 37 (1019C>T) gene polymorphism is associated with subclinical atherosclerosis in women with type 1 and 2 diabetes and in women with central obesity. Physiol Res. 59:1029-1032, 2010.
55. Katakami N, Sakamoto K, Kaneto H, Matsuhisa M, Shimizu I, Ishibashi F, Osonoi T, Kashiwagi A, Kawamori R, Hori M, Yamasaki Y: Association between the connexin37 polymorphism and peripheral arterial disease in subjects with type 2 diabetes. Diabetes Care 32:e53-e54, 2009.
56. Gugl A, Renner W, Seinost G, Brodmann M, Pabst E, Wascher TC, Paulweber B, Iglseder B, Pilger E: Two polymorphisms in the fracalkine receptor CX3CR1 are not associated with peripheral arterial disease. Atherosclerosis 166:339-343, 2003.
57. Renner W, Schallmoser K, Gallippi P, Krauss C, Toplak H, Wascher TC, Pilger E: C242T polymorphism of the p22 phox gene is not associated with peripheral arterial occlusive disease. Atherosclerosis 152:175-179, 2000.

58. Kaur-Knudsen D, Bojesen SE, Nordestgaard BG: Common polymorphisms in CYP2C9, subclinical atherosclerosis and risk of ischemic vascular disease in 52,000 individuals. Pharmacogenomics. J. 9:327-332, 2009.
59. Gretarsdottir S, Baas AF, Thorleifsson G, et al: Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat. Genet 42:692-697, 2010.
60. Eller P, Schgoer W, Mueller T, Tancevski I, Demetz E, Duwensee K, Ritsch A, Haltmayer M, Patsch JR: The K121Q polymorphism of ENPP1 and peripheral arterial disease. Heart Vessels 23:104-107, 2008.
61. Shaker O, Zahra A, Sayed A, Refaat A, El-Khaiat Z, Hegazy G, El-Hindawi K, Ay-El DM: Role of ICAM-1 and E-selectin gene polymorphisms in pathogenesis of PAOD in Egyptian patients. Vasc. Health Risk Manag. 6:9-15, 2010.
62. Renner W, Brodmann M, Pabst E, Stanger O, Wascher TC, Pilger E: The V34L polymorphism of factor XIII and peripheral arterial disease. Int. Angiol. 21:53-57, 2002.
63. Renner W, Koppel H, Brodmann M, Pabst E, Schallmoser K, Toplak H, Wascher TC, Pilger E: Factor II G20210A and factor V G1691A gene mutations and peripheral arterial occlusive disease. Thromb. Haemost. 83:20-22, 2000.
64. Reny JL, Alhenc-Gelas M, Fontana P, Bissery A, Julia PL, Fiessinger JN, Aiach M, Emmerich J: The factor II G20210A gene polymorphism, but not factor V Arg506Gln, is associated with peripheral arterial disease: results of a case-control study. J. Thromb. Haemost. 2:1334-1340, 2004.
65. Sofi F, Lari B, Rogolino A, Marcucci R, Pratesi G, Dorigo W, Pratesi C, Gensini GF, Abbate R, Prisco D: Thrombophilic risk factors for symptomatic peripheral arterial disease. J. Vasc. Surg. 41:255-260, 2005.
66. Mueller T, Marschon R, Dieplinger B, Haidinger D, Gegenhuber A, Poelz W, Webersinke G, Haltmayer M: Factor V Leiden, prothrombin G20210A, and methylenetetrahydrofolate reductase C677T mutations are not associated with chronic limb ischemia: the Linz Peripheral Arterial Disease (LIPAD) study. J. Vasc. Surg. 41:808-815, 2005.
67. Lee AJ, Fowkes FG, Lowe GD, Connor JM, Rumley A: Fibrinogen, factor VII and PAI-1 genotypes and the risk of coronary and peripheral atherosclerosis: Edinburgh Artery Study. Thromb. Haemost. 81:553-560, 1999.
68. Serve E, Reny JL, Akhavan S, Emmerich J, Fischer AM, Tapon-Bretaudiere J: A two adenine insertion polymorphism in the 3' untranslated region of factor VII gene is associated with peripheral arterial disease but not with venous thrombosis. Results of case-control studies. Thromb. Haemost. 98:733-737, 2007.
69. Parry DJ, Grant PJ, Scott DJ: Atherothrombotic risk factor clustering in healthy male relatives of male patients with intermittent claudication. J. Vasc. Surg. 40:891-898, 2004.
70. Fowkes FG, Connor JM, Smith FB, Wood J, Donnan PT, Lowe GD: Fibrinogen genotype and risk of peripheral atherosclerosis. Lancet 339:693-696, 1992.
71. Smith FB, Connor JM, Lee AJ, Cooke A, Lowe GD, Rumley A, Fowkes FG: Relationship of the platelet glycoprotein PlA and fibrinogen T/G+1689 polymorphisms with peripheral arterial disease and ischaemic heart disease. Thromb. Res. 112:209-216, 2003.
72. Klein RL, Hunter SJ, Jenkins AJ, Zheng D, Semler AJ, Clore J, Garvey WT: Fibrinogen is a marker for nephropathy and peripheral vascular disease in type 1 diabetes: studies of plasma fibrinogen and fibrinogen gene polymorphism in the DCCT/EDIC cohort. Diabetes Care 26:1439-1448, 2003.
73. Bahadori B, Uitz E, Dehchamani D, Pilger E, Renner W: The fibrinogen gamma 10034C>T polymorphism is not associated with Peripheral Arterial Disease. Thromb. Res. 126:350-352, 2010.
74. Yazdani-Biuki B, Krippl P, Brickmann K, Fuerst F, Langsenlehner U, Paulweber B, Pilger E, Wascher TC, Brezinschek HP, Renner W: The functional promoter polymorphism of the coagulation factor XII gene is not associated with peripheral arterial disease. Angiology 61:211-215, 2010.

75. Li R, Folsom AR, Sharrett AR, Couper D, Bray M, Tyroler HA: Interaction of the glutathione S-transferase genes and cigarette smoking on risk of lower extremity arterial disease: the Atherosclerosis Risk in Communities (ARIC) study. Atherosclerosis 154:729-738, 2001.
76. Koppel H, Krippl P, Gasser R, Wascher TC, Paulweber B, Pilger E, Renner W: Hemochromatosis gene (HFE) polymorphisms are not associated with peripheral arterial disease. Thromb. Haemost. 91:1258-1259, 2004.
77. Bahadori B, Uitz E, Mayer A, Harauer J, Dam K, Truschnig-Wilders M, Pilger E, Renner W: Polymorphisms of the hypoxia-inducible factor 1 gene and peripheral artery disease. Vasc. Med. 15:371-374, 2010.
78. Delanghe J, Langlois M, Duprez D, De BM, Clement D: Haptoglobin polymorphism and peripheral arterial occlusive disease. Atherosclerosis 145:287-292, 1999.
79. Potaczek DP, Krzanowski M, Sanak M: [Interleukin-1beta C+3953T gene polymorphism and peripheral arterial occlusive disease in Polish population]. Przegl. Lek. 62:830-832, 2005.
80. Libra M, Signorelli SS, Bevelacqua Y, Navolanic PM, Bevelacqua V, Polesel J, Talamini R, Stivala F, Mazzarino MC, Malaponte G: Analysis of G(-174)C IL-6 polymorphism and plasma concentrations of inflammatory markers in patients with type 2 diabetes and peripheral arterial disease. J. Clin. Pathol. 59:211-215, 2006.
81. Renner W, Brodmann M, Winkler M, Washer TC, Pilger E: The PLA1/A2 polymorphism of platelet glycoprotein IIIa is not associated with peripheral arterial disease. Thromb. Haemost. 85:745-746, 2001.
82. Eller P, Schgoer W, Mueller T, Tancevski I, Wehinger A, Ulmer H, Foeger B, Haltmayer M, Ritsch A, Patsch JR: Hepatic lipase polymorphism and increased risk of peripheral arterial disease. J. Intern. Med. 258:344-348, 2005.
83. Dieplinger B, Lingenhel A, Baumgartner N, Poelz W, Dieplinger H, Haltmayer M, Kronenberg F, Mueller T: Increased serum lipoprotein(a) concentrations and low molecular weight phenotypes of apolipoprotein(a) are associated with symptomatic peripheral arterial disease. Clin. Chem. 53:1298-1305, 2007.
84. Catalano M, Cortelazzo A, Yilmaz Y, Perilli E, Carzaniga G, Emanuele E: The LPA gene C93T polymorphism influences plasma lipoprotein(a) levels and is independently associated with susceptibility to peripheral arterial disease. Clin. Chim. Acta 387:109-112, 2008.
85. Khandanpour N, Willis G, Meyer FJ, Armon MP, Loke YK, Wright AJ, Finglas PM, Jennings BA: Peripheral arterial disease and methylenetetrahydrofolate reductase (MTHFR) C677T mutations: A case-control study and meta-analysis. J. Vasc. Surg. 49:711-718, 2009.
86. Schgoer W, Eller P, Mueller T, Tancevski I, Wehinger A, Ulmer H, Sandhofer A, Ritsch A, Haltmayer M, Patsch JR: The MTP -493TT genotype is associated with peripheral arterial disease: results from the Linz Peripheral Arterial Disease (LIPAD) Study. Clin. Biochem. 41:712-716, 2008.
87. Kullo IJ, Greene MT, Boerwinkle E, Chu J, Turner ST, Kardia SL: Association of polymorphisms in NOS3 with the ankle-brachial index in hypertensive adults. Atherosclerosis 196:905-912, 2008.
88. Fontana P, Gaussem P, Aiach M, Fiessinger JN, Emmerich J, Reny JL: P2Y12 H2 haplotype is associated with peripheral arterial disease: a case-control study. Circulation 108:2971-2973, 2003.
89. Unno N, Sakaguchi T, Nakamura T, Yamamoto N, Sugatani J, Miwa M, Konno H: A single nucleotide polymorphism in the plasma PAF acetylhydrolase gene and risk of atherosclerosis in Japanese patients with peripheral artery occlusive disease. J. Surg. Res. 134:36-43, 2006.
90. Catalano M, Cortelazzo A, Santi R, Contino L, Demicheli M, Yilmaz Y, Zorzetto M, Campo I, Lanati N, Emanuele E: The Pro12Ala polymorphism of peroxisome proliferator-activated receptor-gamma2 gene is associated with plasma levels of soluble RAGE (Receptor for Advanced Glycation Endproducts) and the presence of peripheral arterial disease. Clin. Biochem. 41:981-985, 2008.

91. Ritsch A, Sonderegger G, Sandhofer A, Stanzl U, Tancevski I, Eller P, Schgoer W, Wehinger A, Mueller T, Haltmayer M, Patsch JR: Scavenger receptor class B type I polymorphisms and peripheral arterial disease. Metabolism 56:1135-1141, 2007.
92. Jiang YD, Chang YC, Chiu YF, Chang TJ, Li HY, Lin WH, Yuan HY, Chen YT, Chuang LM: SLC2A10 genetic polymorphism predicts development of peripheral arterial disease in patients with type 2 diabetes. SLC2A10 and PAD in type 2 diabetes. BMC Med. Genet 11:126, 2010.
93. Signorelli SS, Barresi V, Musso N, Anzaldi M, Croce E, Fiore V, Condorelli DF: Polymorphisms of steroid 5-alpha-reductase type I (SRD5A1) gene are associated to peripheral arterial disease. J. Endocrinol. Invest 31:1092-1097, 2008.
94. Smadja DM, Loriot MA, Hindorff LA, Mellottee L, Gaussem P, Emmerich J: No clear link between VKORC1 genetic polymorphism and the risk of venous thrombosis or peripheral arterial disease. Thromb. Haemost. 99:970-972, 2008.
95. Stene MC, Frikke-Schmidt R, Nordestgaard BG, Grande P, Schnohr P, Tybjaerg-Hansen A: Functional promoter variant in zinc finger protein 202 predicts severe atherosclerosis and ischemic heart disease. J. Am. Coll. Cardiol. 52:369-377, 2008.
96. Weatherley BD, Nelson JJ, Heiss G, Chambless LE, Sharrett AR, Nieto FJ, Folsom AR, Rosamond WD: The association of the ankle-brachial index with incident coronary heart disease: the Atherosclerosis Risk In Communities (ARIC) study, 1987-2001. BMC Cardiovasc. Disord. 7:3, 2007.
97. Newman AB, Siscovick DS, Manolio TA, Polak J, Fried LP, Borhani NO, Wolfson SK: Ankle-arm index as a marker of atherosclerosis in the Cardiovascular Health Study. Cardiovascular Heart Study (CHS) Collaborative Research Group. Circulation 88:837-845, 1993.
98. Murabito JM, Evans JC, Larson MG, Nieto K, Levy D, Wilson PW: The ankle-brachial index in the elderly and risk of stroke, coronary disease, and death: the Framingham Study. Arch. Intern. Med. 163:1939-1942, 2003.
99. Murabito JM, Evans JC, Nieto K, Larson MG, Levy D, Wilson PW: Prevalence and clinical correlates of peripheral arterial disease in the Framingham Offspring Study. Am. Heart J. 143:961-965, 2002.
100. Kardia SL, Greene MT, Boerwinkle E, Turner ST, Kullo IJ: Investigating the complex genetic architecture of ankle-brachial index, a measure of peripheral arterial disease, in non-Hispanic whites. BMC Med Genomics 1:16, 2008.
101. McDermott MM, Guralnik JM, Albay M, Bandinelli S, Miniati B, Ferrucci L: Impairments of muscles and nerves associated with peripheral arterial disease and their relationship with lower extremity functioning: the InCHIANTI Study. J. Am. Geriatr. Soc. 52:405-410, 2004.
102. Meijer WT, Hoes AW, Rutgers D, Bots ML, Hofman A, Grobbee DE: Peripheral arterial disease in the elderly - The Rotterdam study. Arterioscler. Thromb. Vasc. Biol. 18:185-192, 1998.
103. Hiatt WR, Hoag S, Hamman RF: Effect of diagnostic criteria on the prevalence of peripheral arterial disease. The San Luis Valley Diabetes Study. Circulation 91:1472-1479, 1995.
104. Hirsch AT, Criqui MH, Treat-Jacobson D, Regensteiner JG, Creager MA, Olin JW, Krook SH, Hunninghake DB, Comerota AJ, Walsh ME, McDermott MM, Hiatt WR: Peripheral arterial disease detection, awareness, and treatment in primary care. JAMA 286:1317-1324, 2001.
105. Sadrzadeh Rafie AH, Stefanick ML, Sims ST, Phan T, Higgins M, Gabriel A, Assimes T, Narasimhan B, Nead KT, Myers J, Olin J, Cooke JP: Sex differences in the prevalence of peripheral artery disease in patients undergoing coronary catheterization. Vasc. Med 15:443-450, 2010.

Supplementary Table 6: Overall Meta-analysis Results and Study Specific Results: ABI-SNP Associations with p<10-5

						Meta-analysis results, population isolates excluded					
SNP	Chr	Physical Position	Allele1	Allele2	Allele 1 frequency	Beta	SE	P value	P_{het} value	Total N	Closest gene
rs10757269	9	22062264	a	g	0.5057	0.0056	0.001	2.46E-08	0.2294	35036	CDKN2B
rs9632884	9	22062301	c	g	0.4945	-0.0055	0.001	2.89E-08	0.2555	35035	CDKN2B
rs10811647	9	22055002	c	g	0.5692	0.0054	0.001	8.39E-08	0.3306	35306	CDKN2B
rs10811650	9	22057593	a	g	0.569	0.0054	0.001	9.17E-08	0.323	35305	CDKN2B
rs1537375	9	22106071	t	c	0.4831	0.0053	0.001	1.06E-07	0.05722	35303	CDKN2B
rs2383207	9	22105959	a	g	0.4827	0.0053	0.001	1.12E-07	0.05728	35308	CDKN2B
rs1537374	9	22106046	a	g	0.4829	0.0053	0.001	1.22E-07	0.05755	35303	CDKN2B
rs1004638	9	22105589	a	t	0.4836	0.0053	0.001	1.43E-07	0.06191	35303	CDKN2B
rs944797	9	22105286	t	c	0.484	0.0053	0.001	1.45E-07	0.06293	35303	CDKN2B
rs10511701	9	22102599	t	c	0.4795	0.0053	0.001	1.46E-07	0.07575	35274	CDKN2B
rs7341786	9	22102241	a	c	0.4796	0.0053	0.001	1.46E-07	0.07571	35274	CDKN2B
rs615552	9	22016077	t	c	0.5665	-0.0053	0.001	1.52E-07	0.8319	35397	CDKN2B
rs10738604	9	22015493	a	g	0.3618	-0.0055	0.0011	1.71E-07	0.6335	35364	CDKN2B
rs2383206	9	22105026	a	g	0.4841	0.0052	0.001	1.77E-07	0.07524	35304	CDKN2B
rs10738609	9	22104495	a	g	0.4843	0.0052	0.001	1.79E-07	0.07523	35304	CDKN2B
rs10733376	9	22104469	c	g	0.5156	-0.0052	0.001	1.79E-07	0.0751	35304	CDKN2B
rs10738610	9	22113766	a	c	0.489	0.0053	0.001	1.87E-07	0.0537	35330	CDKN2B
rs1333046	9	22114123	a	t	0.5111	-0.0053	0.001	1.91E-07	0.0535	35330	CDKN2B
rs7859362	9	22095927	t	c	0.4849	0.0052	0.001	2.10E-07	0.07804	35305	CDKN2B
rs1412834	9	22100131	t	c	0.4847	0.0052	0.001	2.10E-07	0.07823	35304	CDKN2B
rs1333042	9	22093813	a	g	0.499	0.0052	0.001	2.15E-07	0.0905	35354	CDKN2B
rs6475606	9	22071850	t	c	0.4873	-0.0051	0.001	2.43E-07	0.229	35411	CDKN2B
rs10116277	9	22071397	t	g	0.4871	-0.0051	0.001	2.80E-07	0.2044	35236	CDKN2B
rs523096	9	22009129	a	g	0.5615	-0.0052	0.001	3.36E-07	0.891	35401	CDKN2B
rs4659996	1	238912747	a	g	0.4846	-0.006	0.0012	4.44E-07	0.3354	28087	GREM2
rs7003385	8	41705907	t	c	0.6656	-0.0053	0.0011	5.24E-07	0.4894	35375	ANK1
rs3218020	9	21987872	a	g	0.3485	-0.0054	0.0011	6.36E-07	0.6896	35095	CDKN2A
rs10811644	9	22015067	a	t	0.542	0.005	0.001	6.63E-07	0.3091	35394	CDKN2B
rs7035484	9	22015240	c	g	0.5421	0.005	0.001	6.86E-07	0.3098	35395	CDKN2B
rs1333043	9	22096731	a	t	0.5036	-0.0049	0.001	8.00E-07	0.04727	35314	CDKN2B
rs564398	9	22019547	t	c	0.582	-0.0049	0.001	9.83E-07	0.8683	35407	CDKN2B
rs634537	9	22022152	t	g	0.5834	-0.0049	0.001	1.01E-06	0.897	35400	CDKN2B
rs7865618	9	22021005	a	g	0.5725	-0.0049	0.001	1.08E-06	0.8914	35411	CDKN2B
rs2069418	9	21999698	c	g	0.542	-0.005	0.001	1.08E-06	0.8796	35224	CDKN2B
rs543830	9	22016639	a	t	0.5849	-0.0049	0.001	1.14E-06	0.8487	35401	CDKN2B
rs2157719	9	22023366	t	c	0.5739	-0.0049	0.001	1.21E-06	0.9112	35382	CDKN2B
rs679038	9	22019080	a	g	0.4156	0.0049	0.001	1.26E-06	0.8549	35409	CDKN2B
rs7044859	9	22008781	a	t	0.4537	-0.0049	0.001	1.36E-06	0.3891	35394	CDKN2B
rs10757272	9	22078260	t	c	0.4904	-0.0048	0.001	1.45E-06	0.06182	35411	CDKN2B
rs1008878	9	22026112	t	g	0.5803	-0.0049	0.001	1.49E-06	0.9102	35196	CDKN2B
rs1333048	9	22115347	a	c	0.4997	0.0049	0.001	1.50E-06	0.03292	35369	CDKN2B
rs10738607	9	22078094	a	g	0.5079	0.0048	0.001	1.60E-06	0.06705	35409	CDKN2B
rs2106120	9	22007101	t	g	0.4538	-0.0048	0.001	1.75E-06	0.3937	35391	CDKN2B
rs2106119	9	22007550	a	g	0.5461	0.0048	0.001	1.76E-06	0.3926	35393	CDKN2B
rs1537370	9	22074310	t	c	0.4894	-0.0048	0.001	1.79E-06	0.1455	35241	CDKN2B
rs10811640	9	22003411	t	g	0.4519	-0.0048	0.001	1.83E-06	0.4239	35350	CDKN2B
rs1556516	9	22090176	c	g	0.4909	-0.0047	0.001	1.91E-06	0.05805	35406	CDKN2B
rs1537371	9	22089568	a	c	0.4909	-0.0047	0.001	1.93E-06	0.05815	35406	CDKN2B
rs1412829	9	22033926	a	g	0.582	-0.0048	0.001	1.95E-06	0.911	35355	CDKN2B
rs10811641	9	22004137	c	g	0.6318	0.0049	0.001	2.14E-06	0.7504	35356	CDKN2B

rs7859727	9	22092165	t	c	0.4892	-0.0047	0.001	2.19E-06	0.05953	35406	CDKN2B
rs1537373	9	22093341	t	g	0.5112	0.0047	0.001	2.28E-06	0.05976	35406	CDKN2B
rs2891168	9	22088619	a	g	0.5095	0.0047	0.001	2.32E-06	0.0655	35410	CDKN2B
rs4977574	9	22088574	a	g	0.5094	0.0047	0.001	2.33E-06	0.06531	35411	CDKN2B
rs2069416	9	22000004	a	t	0.3576	-0.005	0.0011	2.37E-06	0.744	35254	CDKN2B
rs1556515	9	22026367	t	c	0.5815	-0.0048	0.001	2.48E-06	0.9065	35196	CDKN2B
rs1333037	9	22030765	t	c	0.5741	-0.0047	0.001	3.60E-06	0.9059	35356	CDKN2B
rs2099861	8	41700444	t	c	0.3664	-0.0048	0.001	3.64E-06	0.1513	35386	ANK1
rs819750	1	99469651	t	g	0.8772	0.007	0.0015	3.65E-06	0.5103	35278	LPPR4
rs1063192	9	21993367	a	g	0.5676	-0.0047	0.001	3.82E-06	0.9158	35268	CDKN2B
rs819748	1	99467756	a	c	0.8762	0.007	0.0015	4.19E-06	0.5079	35291	LPPR4
rs10757278	9	22114477	a	g	0.5167	0.0047	0.001	4.43E-06	0.0157	35348	CDKN2B
rs1333047	9	22114504	a	t	0.5166	0.0047	0.001	4.47E-06	0.01568	35348	CDKN2B
rs10115049	9	22022119	a	g	0.5233	0.0046	0.001	4.61E-06	0.5317	35383	CDKN2B
rs1360589	9	22035317	t	c	0.5746	-0.0046	0.001	4.62E-06	0.8514	35355	CDKN2B
rs9485528	6	102221473	a	g	0.1699	-0.0061	0.0013	4.63E-06	0.7808	35339	GRIK2
rs1872877	8	41704728	t	c	0.6336	0.0047	0.001	4.66E-06	0.167	35391	ANK1
rs10965212	9	22013795	a	t	0.4768	-0.0046	0.001	4.85E-06	0.3242	35405	CDKN2B
rs819746	1	99466332	a	c	0.8757	0.0069	0.0015	5.04E-06	0.4729	35316	LPPR4
rs4977575	9	22114744	c	g	0.5166	0.0046	0.001	5.47E-06	0.01628	35348	CDKN2B
rs1333049	9	22115503	c	g	0.483	-0.0046	0.001	5.51E-06	0.01224	35358	CDKN2B
rs7030641	9	22044040	t	c	0.573	-0.0046	0.001	5.68E-06	0.8398	35394	CDKN2B
rs2279437	8	41707343	a	g	0.3652	-0.0047	0.001	5.69E-06	0.1507	35368	ANK1
rs819797	1	99459990	a	g	0.866	0.0069	0.0015	5.74E-06	0.3963	35101	LPPR4
rs944801	9	22041670	c	g	0.5739	-0.0046	0.001	5.80E-06	0.8418	35396	CDKN2B
rs722453	7	84037497	a	g	0.5771	0.0054	0.0012	6.43E-06	0.6902	26200	SEMA3A
rs7780912	7	84110129	a	g	0.5769	0.0053	0.0012	6.57E-06	0.7077	26171	SEMA3D
rs7028570	9	22038683	a	g	0.4808	-0.0045	0.001	6.60E-06	0.4655	35342	CDKN2B
rs2151280	9	22024719	a	g	0.4825	-0.0045	0.001	6.73E-06	0.529	35205	CDKN2B
rs1360590	9	22031443	t	c	0.5207	0.0045	0.001	6.81E-06	0.4601	35359	CDKN2B
rs1591136	9	22016834	c	g	0.4767	-0.0045	0.001	6.85E-06	0.3436	35401	CDKN2B
rs10120688	9	22046499	a	g	0.4881	-0.0045	0.001	7.19E-06	0.5001	35408	CDKN2B
rs16824978	2	211380306	t	c	0.2501	-0.0054	0.0012	7.77E-06	0.3732	34950	CPS1
rs7049105	9	22018801	a	g	0.5234	0.0045	0.001	7.88E-06	0.3545	35404	CDKN2B
rs10965224	9	22057276	a	t	0.5993	-0.0045	0.001	8.13E-06	0.5374	35394	CDKN2B
rs10965219	9	22043687	a	g	0.5199	0.0045	0.001	8.19E-06	0.3665	35390	CDKN2B
rs7530667	1	99464320	t	c	0.3874	-0.0045	0.001	8.43E-06	0.1114	35391	LPPR4
rs1537378	9	22051614	a	g	0.3994	0.0045	0.001	8.49E-06	0.6196	35376	CDKN2B
rs8181050	9	22054391	a	g	0.5995	-0.0045	0.001	8.75E-06	0.5883	35405	CDKN2B
rs4977756	9	22058652	a	g	0.6006	-0.0045	0.001	8.82E-06	0.5672	35405	CDKN2B
rs1333039	9	22055657	c	g	0.5995	-0.0045	0.001	8.86E-06	0.5856	35405	CDKN2B
rs7311009	12	94859134	c	g	0.9711	0.0152	0.0034	8.94E-06	0.965	30436	CCDC38
rs10811651	9	22057830	a	g	0.5997	-0.0045	0.001	9.22E-06	0.5821	35404	CDKN2B
rs518394	9	22009673	c	g	0.4011	0.0047	0.0011	9.38E-06	0.8702	35252	CDKN2B
rs819756	1	99474518	t	c	0.8961	0.0074	0.0017	9.71E-06	0.5301	35257	LPPR4
rs2241896	8	41674630	t	c	0.6369	0.0047	0.0011	9.99E-06	0.1807	35243	ANK1

Atherosclerosis Risk in Communities Study

af	oevar_aric	coded_aric	noncoded_aric	beta_aric	se_aric	p_aric	n_aric	maf_aric
0.4879	0.9945	G	A	-0.00183463	0.00200732	0.36068207	7630	0.4879
0.488	0.9955	C	G	-0.00183487	0.00200629	0.36037386	7630	0.488
0.5733	0.9732	C	G	0.00319315	0.0020569	0.12054388	7630	0.4267
0.5732	0.9729	A	G	0.00318621	0.00205763	0.12148582	7630	0.4268
0.5128	0.9759	C	T	-5.76E-05	0.00203077	0.97735436	7630	0.4872
0.5124	0.9753	G	A	-4.98E-05	0.00203108	0.98043457	7630	0.4876
0.5126	0.9756	G	A	-5.31E-05	0.00203097	0.97913034	7630	0.4874
0.5123	0.9753	T	A	-4.81E-05	0.00203109	0.98108634	7630	0.4877
0.5122	0.9752	C	T	-4.64E-05	0.00203108	0.98176423	7630	0.4878
0.5144	0.9656	C	T	-8.89E-05	0.00203988	0.96523633	7630	0.4856
0.5143	0.9656	C	A	-8.82E-05	0.00203976	0.96551416	7630	0.4857
0.5582	0.9978	T	C	-0.0031378	0.0020047	0.11751214	7630	0.4418
0.3568	0.9858	A	G	-0.00397774	0.00210414	0.05869572	7630	0.3568
0.5121	0.9752	G	A	-4.42E-05	0.00203104	0.98264733	7630	0.4879
0.512	0.9752	G	A	-4.26E-05	0.002031	0.98327579	7630	0.488
0.5119	0.9752	C	G	-4.00E-05	0.00203085	0.98427083	7630	0.4881
0.5086	0.9879	C	A	-0.00043888	0.00202089	0.82805345	7630	0.4914
0.5087	0.9881	A	T	-0.00043936	0.00202062	0.82784477	7630	0.4913
0.5112	0.9757	C	T	-2.75E-05	0.00202968	0.98918975	7630	0.4888
0.5114	0.9754	C	T	-3.22E-05	0.00203018	0.98734162	7630	0.4886
0.5007	0.9932	G	A	-2.91E-06	0.00201531	0.99884963	7630	0.4993
0.4811	0.9985	T	C	-0.00101356	0.0019989	0.6120711	7630	0.4811
0.4809	0.9934	T	G	-0.00105318	0.0020046	0.59927153	7630	0.4809
0.5536	0.9997	A	G	-0.0026719	0.00200147	0.18185367	7630	0.4464
0.5166	0.8771	G	A	0.00185897	0.00212836	0.38237946	7630	0.4834
0.6666	0.9984	T	C	-0.00504406	0.00212636	0.01769082	7630	0.3334
0.3409	0.9395	A	G	-0.00326476	0.00218813	0.13566875	7630	0.3409
0.4512	0.9981	T	A	-0.00254969	0.00201594	0.20592022	7630	0.4512
0.4512	0.998	G	C	-0.00253295	0.00201604	0.20893363	7630	0.4512
0.4976	0.9816	A	T	0.00023452	0.00201663	0.90740828	7630	0.4976
0.5756	0.9997	T	C	-0.0028059	0.00201974	0.16473219	7630	0.4244
0.5753	0.9967	T	G	-0.00288453	0.00202053	0.15337896	7630	0.4247
0.5668	0.9979	A	G	-0.00259143	0.00201473	0.19832451	7630	0.4332
0.5346	0.9718	C	G	-0.00196773	0.00203216	0.33284865	7630	0.4654
0.5772	0.9994	A	T	-0.00274235	0.00201929	0.17440808	7630	0.4228
0.567	0.9961	T	C	-0.00256224	0.00201604	0.20372039	7630	0.433
0.5767	0.9983	G	A	-0.0027566	0.00202022	0.17237802	7630	0.4233
0.4466	0.9996	A	T	-0.00209805	0.00201716	0.29824611	7630	0.4466
0.4848	0.9977	T	C	0.00035805	0.00200593	0.85831437	7630	0.4848
0.5762	0.9661	T	G	-0.00266527	0.00204774	0.19303248	7630	0.4238
0.4958	0.999	C	A	-0.00018765	0.00200289	0.92534493	7630	0.4958
0.4847	0.9929	G	A	0.00030927	0.00201038	0.87772422	7630	0.4847
0.4464	0.9984	T	G	-0.00210092	0.00201877	0.29796941	7630	0.4464

0.4465	0.9989	G	A	-0.00209888	0.0020181	0.29828183	7630	0.4465
0.4837	0.9696	T	C	-2.58E-05	0.00203274	0.98988367	7630	0.4837
0.4442	0.9892	T	G	-0.00213926	0.00203169	0.29232119	7630	0.4442
0.4866	0.9995	C	G	0.00036765	0.00200258	0.85431714	7630	0.4866
0.4865	0.9998	A	C	0.00037253	0.00200233	0.8523868	7630	0.4865
0.5731	0.9879	A	G	-0.00263971	0.00202643	0.19266374	7630	0.4269
0.3587	0.9889	G	C	-0.00271476	0.00210413	0.19694572	7630	0.3587
0.4866	0.9991	T	C	0.00036118	0.00200292	0.85687916	7630	0.4866
0.4866	0.9991	G	T	0.00035985	0.00200299	0.85740336	7630	0.4866
0.4865	0.9999	G	A	0.00037458	0.00200223	0.85157869	7630	0.4865
0.4865	0.9999	G	A	0.00037366	0.00200222	0.85193679	7630	0.4865
0.3504	0.9757	A	T	-0.00239867	0.00213083	0.26025124	7630	0.3504
0.5765	0.9628	T	C	-0.00252026	0.00205112	0.21913909	7630	0.4235
0.5662	0.989	T	C	-0.00229685	0.0020215	0.25582692	7630	0.4338
0.367	0.9975	T	C	-0.00470226	0.00206081	0.02250919	7630	0.367
0.8766	0.9602	T	G	0.00949865	0.00306533	0.00194718	7630	0.1234
0.5602	0.9858	A	G	-0.00194817	0.00203368	0.33803723	7630	0.4398
0.8749	0.973	A	C	0.00925206	0.00302879	0.0022568	7630	0.1251
0.4795	0.9875	G	A	0.00053336	0.00201649	0.79137114	7630	0.4795
0.4797	0.9882	T	A	0.00053025	0.00201571	0.79247952	7630	0.4797
0.4738	0.9959	G	A	-0.00256005	0.00201271	0.20335971	7630	0.4738
0.566	0.9955	T	C	-0.0014943	0.00201459	0.45819596	7630	0.434
0.8405	0.9736	G	A	0.00598268	0.00269535	0.02644875	7630	0.1595
0.3678	0.9982	C	T	-0.00474185	0.00205868	0.02126491	7630	0.3678
0.4701	0.9998	A	T	-0.00212833	0.00200796	0.28912529	7630	0.4701
0.8746	0.9763	A	C	0.00920232	0.00302111	0.00232309	7630	0.1254
0.4798	0.9889	G	C	0.00052789	0.00201488	0.79329864	7630	0.4798
0.4796	0.9948	C	G	0.00062719	0.00201003	0.75498597	7630	0.4796
0.5652	0.9967	T	C	-0.0014528	0.00201292	0.4704035	7630	0.4348
0.3676	0.9979	A	G	-0.00484375	0.00205939	0.01867719	7630	0.3676
0.8624	0.8951	A	G	0.00868779	0.00303747	0.00423886	7630	0.1376
0.5657	0.9975	C	G	-0.00145724	0.00201248	0.46895425	7630	0.4343

Cardiovascular Health Study

af_chs	oevar_chs	coded_chs	noncoded_chs	beta_chs	se_chs	p_chs	n_chs	maf_chs
0.47619073	0.93163899	A	G	0.00688849	0.00339263	0.04231299	3193	0.47619073
0.47615143	0.93186249	G	C	0.00688602	0.0033922	0.04236097	3193	0.47615143
0.43995866	0.98017303	G	C	-0.00460375	0.00334966	0.16932005	3193	0.43995866
0.44000266	0.98019466	G	A	-0.00460657	0.0033496	0.16905017	3193	0.44000266
0.46460179	0.98269811	T	C	0.00543526	0.00334237	0.10391326	3193	0.46460179
0.46799092	1.00475623	A	G	0.00535525	0.00330539	0.10519916	3193	0.46799092
0.46724225	0.99264554	A	G	0.00538876	0.003325	0.1050862	3193	0.46724225
0.47253868	0.98847442	A	T	0.0052339	0.00333005	0.11601603	3193	0.47253868
0.47552067	0.98341614	T	C	0.00521381	0.0033371	0.11819953	3193	0.47552067
0.45960492	0.91913763	T	C	0.00549199	0.00344926	0.11133459	3193	0.45960492
0.45958284	0.91844525	A	C	0.00549058	0.00345039	0.11154356	3193	0.45958284
0.4140642	0.85768652	C	T	0.00606861	0.00360725	0.09250298	3193	0.4140642
0.34580629	0.73440757	A	G	-0.00664051	0.00404949	0.10103832	3193	0.34580629
0.47610085	0.98181345	A	G	0.00517855	0.00333915	0.1209355	3193	0.47610085
0.47706044	0.97938435	A	G	0.00510229	0.00334188	0.12681833	3193	0.47706044
0.47758926	0.97805503	G	C	0.00509845	0.00334373	0.12731361	3193	0.47758926
0.47062026	0.90384121	A	C	0.00558813	0.00346506	0.10680841	3193	0.47062026
0.46966708	0.90054547	T	A	0.00557733	0.00347125	0.10811526	3193	0.46966708
0.47518822	0.95985821	T	C	0.0052924	0.0033725	0.11658255	3193	0.47518822
0.47459395	0.96081271	T	C	0.005273	0.00337124	0.11779056	3193	0.47459395
0.52002114	0.95431687	A	G	0.00592529	0.0033536	0.07725448	3193	0.479979
0.5013705	0.99732119	C	T	0.00657617	0.00325919	0.04361916	3193	0.49863
0.49939712	1.01003772	T	G	-0.00670693	0.00323948	0.03841743	3193	0.49939712
0.42033432	0.86356124	G	A	0.00622611	0.00359893	0.08363193	3193	0.42033432
0.48953696	0.97817916	A	G	-0.0071575	0.00341608	0.03614992	3193	0.48953696
0.35147087	0.74960305	C	T	0.00556292	0.0040051	0.16484488	3193	0.35147087
0.36293846	0.98399398	A	G	-0.00639755	0.00345063	0.06373542	3193	0.36293846
0.52303304	0.84457689	A	T	0.0074032	0.00361997	0.04084512	3193	0.476967
0.52451769	0.84425702	C	G	0.00735993	0.0036214	0.04211865	3193	0.475482
0.502443	0.87453154	T	A	0.00533974	0.00351843	0.1291033	3193	0.497557
0.4226746	0.91389175	C	T	0.00521765	0.00350143	0.13618504	3193	0.4226746
0.40220952	0.97376389	G	T	0.00520125	0.00342372	0.12871714	3193	0.40220952
0.43391575	0.92426587	G	A	0.00545496	0.00346878	0.11581415	3193	0.43391575
0.45653727	0.86946078	G	C	0.00643676	0.00356576	0.07104975	3193	0.45653727
0.39945412	0.88399562	T	A	0.00587369	0.00357222	0.10012097	3193	0.39945412
0.41525321	0.98792437	C	T	0.0053248	0.00336926	0.11401375	3193	0.41525321
0.40445803	0.90029984	A	G	0.00552049	0.00354139	0.11903271	3193	0.40445803
0.47357172	0.86093229	A	T	-0.0076699	0.00358862	0.03257474	3193	0.47357172
0.50445725	0.96349096	C	T	0.00636697	0.00331902	0.0550697	3193	0.495543
0.41612606	0.97776287	G	T	0.00546894	0.00338701	0.10638036	3193	0.41612606
0.49081992	0.79874005	A	C	0.00557985	0.0036674	0.12814182	3193	0.49081992
0.50287363	0.96630408	A	G	0.00638104	0.00331304	0.05409942	3193	0.497126
0.52554244	0.86140044	G	T	0.00767844	0.00358726	0.03231609	3193	0.474458

0.52554322	0.86139719	A	G	0.00767829	0.00358726	0.03231992	3193	0.474457
0.50023223	0.96879241	C	T	0.00647096	0.00330849	0.05048141	3193	0.499768
0.46146774	0.90708127	T	G	-0.00725322	0.00351486	0.03905719	3193	0.46146774
0.48435343	0.89305765	C	G	-0.00641169	0.00345436	0.06343714	3193	0.48435343
0.48516411	0.89443313	A	C	-0.00638951	0.00345079	0.06408182	3193	0.48516411
0.40386893	0.99322843	G	A	0.00573334	0.00340083	0.09182257	3193	0.40386893
0.38296821	0.8195557	G	C	-0.00611346	0.00374023	0.10215111	3193	0.38296821
0.53625744	0.90208384	C	T	0.00602395	0.00344272	0.08015857	3193	0.463743
0.4581508	0.90751773	G	T	-0.00590263	0.00343653	0.08586657	3193	0.4581508
0.51939211	0.90910282	A	G	0.00636452	0.00342129	0.06284728	3193	0.480608
0.51801676	0.9105468	A	G	0.00639234	0.00341783	0.06144295	3193	0.481983
0.35873974	0.82350797	A	T	-0.00629036	0.00378829	0.09681871	3193	0.35873974
0.40686988	0.95479807	C	T	0.00555421	0.00343211	0.10559598	3193	0.40686988
0.40867836	0.95718608	C	T	0.00564112	0.0034349	0.10052877	3193	0.40867836
0.34295161	0.7625445	T	C	-0.00667154	0.00404341	0.09894737	3193	0.34295161
0.11465878	1.02855338	G	T	-0.0089047	0.00510546	0.08113237	3193	0.11465878
0.42177889	0.9854606	G	A	0.00618952	0.00337143	0.06637631	3193	0.42177889
0.11589821	0.94777648	C	A	-0.00950566	0.00528459	0.07205815	3193	0.11589821
0.47202897	0.73894179	G	A	-0.00432705	0.00378431	0.25286432	3193	0.47202897
0.47203758	0.73891842	T	A	-0.00432702	0.00378434	0.25287169	3193	0.47203758
0.51234028	0.96262635	A	G	0.00591505	0.00339423	0.08139019	3193	0.48766
0.40722127	0.91984761	C	T	0.00551321	0.00350048	0.11525955	3193	0.40722127
0.17786439	0.72694677	A	G	-0.00636832	0.00525899	0.22591838	3193	0.17786439
0.34022675	0.74826656	C	T	-0.0068743	0.00408413	0.09234099	3193	0.34022675
0.50601049	0.87266629	T	A	0.0071088	0.00354895	0.04516951	3193	0.49399
0.11523505	0.93153156	C	A	-0.00955237	0.00534274	0.07378982	3193	0.11523505
0.47204995	0.73888869	G	C	-0.00432702	0.00378444	0.25288402	3193	0.47204995
0.47208988	0.73879111	C	G	-0.00432851	0.00378452	0.25273098	3193	0.47208988
0.41445944	0.90096971	C	T	0.00571738	0.00352198	0.10451569	3193	0.41445944
0.33333636	0.74162621	A	G	-0.00729095	0.00412523	0.07716027	3193	0.33333636
0.12394159	0.81307176	G	A	-0.0105103	0.00554024	0.05781687	3193	0.12394159
0.40734591	0.91841393	G	C	0.00552373	0.00350138	0.11466036	3193	0.40734591
0.43618212	0.98678838	G	A	-0.00997384	0.00335105	0.00291716	3193	0.43618212
0.42606389	0.98781614	G	A	-0.00957766	0.00335615	0.00432044	3193	0.42606389
0.50458801	0.93410068	G	A	0.00616011	0.00344202	0.07350556	3193	0.495412
0.48017444	0.99662926	A	G	-0.0052702	0.00335759	0.11649936	3193	0.48017444
0.50501206	0.95617204	T	C	0.00602889	0.00340763	0.07685549	3193	0.494988
0.50532524	0.88919568	G	C	0.00660862	0.00351643	0.06019612	3193	0.494675
0.49609756	1.00802829	A	G	-0.00505888	0.00332674	0.12834269	3193	0.49609756
0.28297714	0.76069416	T	C	-0.01211976	0.00424259	0.00428081	3193	0.28297714
0.50745803	0.90132606	A	G	0.00634113	0.00349897	0.06994204	3193	0.492542
0.39341325	0.97036766	T	A	0.00656648	0.00340892	0.05407159	3193	0.39341325
0.50413514	0.90046898	A	G	0.00611493	0.00350442	0.08099882	3193	0.495865
0.37948309	0.93475383	T	C	-0.00857897	0.00355796	0.01589985	3193	0.37948309
0.38823974	0.95621341	A	G	0.00667086	0.0034426	0.05265475	3193	0.38823974
0.39333182	0.9697676	G	A	0.00656451	0.00341011	0.05422789	3193	0.39333182
0.38758268	0.99392834	G	A	0.00641695	0.0033747	0.05723811	3193	0.38758268

0.39341967	0.97034506	G	C	0.00656557	0.00340895	0.05410666	3193	0.39341967
0.03742452	0.1798872	G	C	-0.01350262	0.02471886	0.58489552	3193	0.03742452
0.39230817	0.97415835	G	A	0.00655946	0.0034045	0.05401669	3193	0.39230817
0.36855105	0.65881478	C	G	0.00739739	0.0042007	0.07823982	3193	0.36855105
0.08911697	0.57027126	C	T	-0.01536755	0.00764962	0.04454455	3193	0.08911697
0.34480066	0.64922581	C	T	-0.00307645	0.00445638	0.48997656	3193	0.34480066

Rotterdam Study II

af_eplus	oevar_eplus	coded_eplus	noncoded_eplus	beta_eplus	se_eplus	p_eplus	n_eplus	maf_eplus
0.465425	0.9828	G	A	-0.0145255	0.00469681	0.00200146	1642	0.465425
0.465443	0.9829	C	G	-0.0145092	0.00469633	0.00202275	1642	0.465443
0.583946	0.9933	C	G	0.0151418	0.00473399	0.0013959	1642	0.416054
0.583789	0.9933	A	G	0.0151214	0.00473386	0.00141641	1642	0.416211
0.482092	0.994	C	T	-0.01482	0.00463816	0.00141189	1642	0.482092
0.484354	0.9996	G	A	-0.0150682	0.00462101	0.00112386	1642	0.484354
0.483223	0.9961	G	A	-0.0149556	0.00463135	0.00125505	1642	0.483223
0.483833	0.9985	T	A	-0.0150147	0.00462604	0.00118493	1642	0.483833
0.483441	0.9989	C	T	-0.0149565	0.00462706	0.00124119	1642	0.483441
0.488976	0.9781	C	T	-0.0152647	0.00467531	0.00110763	1642	0.488976
0.488942	0.9779	C	A	-0.0152631	0.00467574	0.00111015	1642	0.488942
0.562711	0.9727	T	C	-0.00452235	0.00475321	0.34115745	1642	0.437289
0.358371	0.9563	A	G	-0.0167126	0.00499561	0.00083238	1642	0.358371
0.483405	0.9987	G	A	-0.0149551	0.00462743	0.00124368	1642	0.483405
0.483357	0.9985	G	A	-0.0149514	0.00462796	0.00124867	1642	0.483357
0.483321	0.9983	C	G	-0.0149502	0.00462836	0.00125111	1642	0.483321
0.475855	0.9559	C	A	-0.0149356	0.00474763	0.00167174	1642	0.475855
0.475823	0.953	A	T	-0.0149587	0.00475489	0.00167149	1642	0.475823
0.482878	0.9959	C	T	-0.0149226	0.00463349	0.0012932	1642	0.482878
0.482953	0.9963	C	T	-0.0149292	0.00463259	0.00128406	1642	0.482953
0.469639	0.983	G	A	-0.0134829	0.00469501	0.00410598	1642	0.469639
0.458942	0.9989	T	C	-0.0122774	0.00467777	0.00870369	1642	0.458942
0.458943	0.9991	T	G	-0.0122756	0.00467739	0.00870807	1642	0.458943
0.556484	0.9707	A	G	-0.00409128	0.0047576	0.38958305	1642	0.443516
0.530285	0.9929	G	A	0.00485961	0.0047631	0.30738464	1642	0.469715
0.653729	0.9999	T	C	-0.00665627	0.00485232	0.16999811	1642	0.346271
0.336983	0.9215	A	G	-0.0175588	0.00518962	0.00072583	1642	0.336983
0.455511	0.9723	T	A	-0.0118925	0.00475131	0.01234482	1642	0.455511
0.455473	0.9725	G	C	-0.0118913	0.0047507	0.01234259	1642	0.455473
0.47866	0.9857	A	T	-0.0148727	0.00466037	0.00143101	1642	0.47866
0.57066	0.9946	T	C	-0.00465588	0.00470929	0.32260774	1642	0.42934
0.568938	0.9986	T	G	-0.00451543	0.00469878	0.3363367	1642	0.431062
0.558149	0.9948	A	G	-0.00608508	0.00467028	0.19244106	1642	0.441851
0.530856	0.9599	C	G	-0.00537446	0.00473921	0.25658122	1642	0.469144
0.572089	0.9977	A	T	-0.00477581	0.00470359	0.30971986	1642	0.427911
0.556438	0.9993	T	C	-0.0059306	0.00465881	0.20285909	1642	0.443562
0.572059	0.9977	G	A	-0.00476572	0.00470365	0.31074657	1642	0.427941
0.450091	0.968	A	T	-0.0114994	0.0047719	0.01598912	1642	0.450091
0.459536	0.9859	T	C	-0.0125143	0.00470698	0.00787396	1642	0.459536
0.556497	0.9988	T	G	-0.00591184	0.00466016	0.20442294	1642	0.443503
0.471586	0.931	C	A	-0.0149792	0.00481525	0.00188288	1642	0.471586
0.459504	0.9865	G	A	-0.0125027	0.0047056	0.00791348	1642	0.459504
0.450315	0.9673	T	G	-0.0114986	0.00477362	0.01603456	1642	0.450315

0.450278	0.9674	G	A	-0.0114997	0.00477335	0.01601861	1642	0.450278
0.458972	0.9979	T	C	-0.0122962	0.00468025	0.00863719	1642	0.458972
0.450465	0.9667	T	G	-0.0114816	0.00477528	0.0162291	1642	0.450465
0.465009	0.9917	C	G	-0.0131286	0.00468252	0.00507696	1642	0.465009
0.464947	0.992	A	C	-0.0131195	0.00468221	0.00510461	1642	0.464947
0.571009	0.9978	A	G	-0.00423072	0.00471136	0.36896153	1642	0.428991
0.363925	0.9618	G	C	-0.014111	0.00498268	0.00465069	1642	0.363925
0.465145	0.9912	T	C	-0.0131493	0.00468318	0.00501437	1642	0.465145
0.465254	0.9908	G	T	-0.0131668	0.00468369	0.00496144	1642	0.465254
0.464897	0.9922	G	A	-0.0131114	0.00468196	0.00512963	1642	0.464897
0.464858	0.9923	G	A	-0.013106	0.00468177	0.00514606	1642	0.464858
0.351306	0.9533	A	T	-0.014129	0.00504732	0.00514723	1642	0.351306
0.556976	0.9967	T	C	-0.00598639	0.00466715	0.19944825	1642	0.443024
0.558446	0.9965	T	C	-0.00579713	0.00467405	0.21469969	1642	0.441554
0.369241	0.9994	T	C	0.00606699	0.00477462	0.20367981	1642	0.369241
0.872412	1	T	G	-0.00220322	0.00704311	0.75427654	1642	0.127588
0.545683	0.999	A	G	-0.00620274	0.00466538	0.1835257	1642	0.454317
0.872411	1	A	C	-0.0022026	0.0070431	0.75434293	1642	0.127589
0.459954	0.9132	G	A	-0.015191	0.00487398	0.00184532	1642	0.459954
0.459933	0.9118	T	A	-0.0152026	0.00487797	0.00184645	1642	0.459933
0.461189	0.9992	G	A	-0.0116514	0.00468266	0.01286864	1642	0.461189
0.558388	0.9898	T	C	-0.0055904	0.00469955	0.23403483	1642	0.441612
0.83711	0.9999	G	A	0.00311825	0.00626805	0.61864756	1642	0.16289
0.369112	0.9998	C	T	0.00598818	0.00477357	0.2095113	1642	0.369112
0.46493	0.9967	A	T	-0.0118836	0.00467944	0.01112989	1642	0.46493
0.87241	1	A	C	-0.00220252	0.00704308	0.75435155	1642	0.12759
0.459921	0.911	G	C	-0.0152138	0.00488016	0.00184084	1642	0.459921
0.45983	0.9043	C	G	-0.0152518	0.00489902	0.00186741	1642	0.45983
0.558332	0.9894	T	C	-0.00552529	0.00469984	0.23955342	1642	0.441668
0.368538	0.997	A	G	0.00593791	0.00478094	0.21406642	1642	0.368538
0.862253	0.9224	A	G	-0.00135591	0.00708063	0.84804708	1642	0.137747
0.558367	0.9897	C	G	-0.00556553	0.00469972	0.23613843	1642	0.441633
0.418647	0.9999	G	A	0.00404727	0.0048044	0.399323	1642	0.418647
0.419305	1	G	A	0.00415108	0.00480224	0.38713081	1642	0.419305
0.467556	0.9906	A	G	-0.01103	0.00470517	0.01909341	1642	0.467556
0.461122	0.9995	A	G	-0.0116321	0.00468196	0.01300507	1642	0.461122
0.464256	0.9946	C	T	-0.0115853	0.00468894	0.01351161	1642	0.464256
0.464661	0.9984	C	G	-0.0118743	0.00467568	0.01112838	1642	0.464661
0.4752	0.9962	A	G	-0.00977621	0.00467291	0.03643988	1642	0.4752
0.75216	0.9971	C	T	0.0138221	0.00523834	0.00835302	1642	0.24784
0.464589	0.9989	G	A	-0.0118709	0.00467452	0.0111312	1642	0.464589
0.416947	0.9971	T	A	0.00455732	0.00474609	0.3367152	1642	0.416947
0.463615	0.9791	G	A	-0.0120559	0.00472348	0.01073027	1642	0.463615
0.593284	0.9939	C	T	0.00897128	0.00475543	0.05920557	1642	0.406716
0.583497	0.9934	G	A	-0.00503299	0.00475273	0.28940226	1642	0.416503
0.417108	0.9962	G	A	0.00457507	0.00474759	0.3349895	1642	0.417108
0.416744	0.9981	G	A	0.00453523	0.00474416	0.33886542	1642	0.416744

0.416987	0.9968	G	C	0.00456153	0.00474646	0.33630586	1642	0.416987
0.967328	0.8988	C	G	0.0101079	0.0140318	0.47107084	1642	0.032672
0.416772	0.998	G	A	0.00453922	0.00474445	0.3384697	1642	0.416772
0.607156	0.8204	G	C	-0.00289013	0.00524028	0.58106319	1642	0.392844
0.892022	0.9696	T	C	-0.00376301	0.00768726	0.62427979	1642	0.107978
0.362713	0.8209	C	T	0.00105613	0.00537832	0.84422909	1642	0.362713

Family Heart Study

af_famhs	oevar_famhs	coded_famhs	oncoded_famh	beta_famhs	se_famhs	p_famhs	n_famhs	maf_famhs
0.5144	0.9627	A	G	0.006716	0.004168	0.10736016	1736	0.4856
0.4856	0.9632	C	G	-0.0067	0.004168	0.10820111	1736	0.4856
0.5798	0.9946	C	G	0.007404	0.004133	0.07347037	1736	0.4202
0.5798	0.995	A	G	0.007415	0.004132	0.07296254	1736	0.4202
0.5244	0.9897	C	T	-0.00803	0.004109	0.05101198	1736	0.4756
0.4708	0.9992	A	G	0.00798	0.004099	0.05178223	1736	0.4708
0.4737	0.9914	A	G	0.008021	0.004108	0.05112133	1736	0.4737
0.4733	0.9967	A	T	0.007914	0.004095	0.05351406	1736	0.4733
0.526	0.9988	C	T	-0.00788	0.004089	0.05422653	1736	0.474
0.5357	0.9625	C	T	-0.00779	0.004193	0.06338047	1736	0.4643
0.4643	0.9622	A	C	0.007795	0.004194	0.06328012	1736	0.4643
0.4399	0.9899	C	T	0.01009	0.004136	0.01485334	1736	0.4399
0.3523	0.995	A	G	-0.0096	0.004234	0.02349133	1736	0.3523
0.4741	0.9983	A	G	0.007884	0.00409	0.05411029	1736	0.4741
0.4741	0.9978	A	G	0.007887	0.004091	0.05408672	1736	0.4741
0.5258	0.9976	C	G	-0.00789	0.004091	0.05404816	1736	0.4742
0.4837	0.9652	A	C	0.007377	0.004152	0.07581954	1736	0.4837
0.5163	0.964	A	T	-0.00739	0.004154	0.07564577	1736	0.4837
0.5246	0.9909	C	T	-0.00795	0.004105	0.0529997	1736	0.4754
0.5248	0.9921	C	T	-0.00794	0.004103	0.05314592	1736	0.4752
0.4945	0.9742	A	G	0.007049	0.004121	0.08742494	1736	0.4945
0.5177	0.9978	C	T	0.005345	0.004087	0.1911306	1736	0.4823
0.5173	0.9977	G	T	0.005314	0.004088	0.19383488	1736	0.4827
0.5545	0.9907	A	G	-0.01018	0.004153	0.01431974	1736	0.4455
0.4909	0.996	A	G	-0.00367	0.004133	0.37461227	1736	0.4909
0.3295	0.9994	C	T	0.000352	0.004169	0.93264816	1736	0.3295
0.3462	0.9975	A	G	-0.01048	0.004258	0.01394803	1736	0.3462
0.5521	0.9923	A	T	0.007497	0.004136	0.07009121	1736	0.4479
0.5521	0.9926	C	G	0.00749	0.004135	0.07032631	1736	0.4479
0.518	0.9771	A	T	-0.00805	0.004133	0.05173598	1736	0.482
0.4184	0.9983	C	T	0.008942	0.004093	0.02908529	1736	0.4184
0.4183	0.9988	G	T	0.008923	0.004092	0.02942418	1736	0.4183
0.5708	0.999	A	G	-0.00814	0.004106	0.04779036	1736	0.4292
0.5346	0.9795	C	G	-0.00896	0.004158	0.03128112	1736	0.4654
0.5818	0.9968	A	T	-0.00917	0.004108	0.02581541	1736	0.4182
0.4292	0.9996	C	T	0.008114	0.004106	0.04836752	1736	0.4292
0.4182	0.9972	A	G	0.009152	0.004107	0.02605562	1736	0.4182
0.4439	0.9912	A	T	-0.00767	0.004166	0.06591506	1736	0.4439
0.5045	0.9947	C	T	0.006845	0.004088	0.09430636	1736	0.4955
0.429	0.9987	G	T	0.008179	0.004111	0.0468475	1736	0.429
0.4912	0.9207	A	C	0.00777	0.004257	0.06818866	1736	0.4912
0.5036	0.9937	A	G	0.007157	0.004084	0.07996042	1736	0.4964
0.5561	0.9931	G	T	0.007499	0.004157	0.07150495	1736	0.4439
0.5562	0.9935	A	G	0.007481	0.004157	0.07219413	1736	0.4438

0.5183	0.9988	C	T	0.005405	0.004084	0.18598291	1736	0.4817
0.5559	0.9919	G	T	0.007591	0.004156	0.06802409	1736	0.4441
0.4967	0.9937	C	G	-0.00693	0.004083	0.08969196	1736	0.4967
0.4966	0.9941	A	C	-0.00693	0.004082	0.08974031	1736	0.4966
0.5834	0.9968	A	G	-0.00868	0.004094	0.03411268	1736	0.4166
0.6441	0.9949	C	G	0.01053	0.004256	0.01344365	1736	0.3559
0.5031	0.9923	C	T	0.006938	0.004085	0.08968342	1736	0.4969
0.497	0.992	G	T	-0.00694	0.004086	0.08966676	1736	0.497
0.5035	0.9944	A	G	0.006928	0.004081	0.08985034	1736	0.4965
0.5036	0.995	A	G	0.006926	0.00408	0.08984471	1736	0.4964
0.3504	0.9739	A	T	-0.01078	0.004311	0.01255035	1736	0.3504
0.428	0.9925	C	T	0.008047	0.004121	0.05106787	1736	0.428
0.4278	0.995	C	T	0.007852	0.004106	0.05603583	1736	0.4278
0.6097	0.9989	C	T	0.003148	0.004133	0.4462965	1736	0.3903
0.1268	1	G	T	-0.00717	0.006042	0.23524773	1736	0.1268
0.5629	0.9986	A	G	-0.00814	0.004115	0.04819188	1736	0.4371
0.8732	1	A	C	0.007175	0.006042	0.23524773	1736	0.1268
0.497	0.9359	A	G	0.007969	0.004209	0.05856344	1736	0.497
0.4971	0.9334	A	T	0.007998	0.004215	0.05802156	1736	0.4971
0.5344	0.9983	A	G	0.00575	0.004102	0.16123042	1736	0.4656
0.4313	0.9894	C	T	0.007726	0.004104	0.05997402	1736	0.4313
0.1673	0.9998	A	G	-0.00743	0.005495	0.17655057	1736	0.1673
0.3906	0.9992	C	T	-0.00314	0.004131	0.44696475	1736	0.3906
0.4702	0.9948	A	T	-0.00665	0.004103	0.10553328	1736	0.4702
0.8732	1	A	C	0.007175	0.006042	0.23526788	1736	0.1268
0.4981	0.9173	C	G	0.008188	0.004254	0.05447115	1736	0.4981
0.5011	0.905	C	G	-0.00834	0.004284	0.05178666	1736	0.4989
0.4319	0.9887	C	T	0.007678	0.004103	0.06154363	1736	0.4319
0.387	0.9955	A	G	-0.00257	0.004152	0.53600811	1736	0.387
0.8627	0.924	A	G	0.004546	0.006058	0.45314912	1736	0.1373
0.5683	0.9897	C	G	-0.00771	0.004102	0.06047336	1736	0.4317
0.5808	0.9997	A	G	0.004295	0.004122	0.29764094	1736	0.4192
0.5801	0.9997	A	G	0.004789	0.004119	0.24517144	1736	0.4199
0.4725	0.9907	A	G	-0.00676	0.004104	0.09992396	1736	0.4725
0.4657	0.9978	A	G	-0.00581	0.004105	0.15745608	1736	0.4657
0.475	0.9918	C	T	-0.00632	0.004098	0.12358667	1736	0.475
0.469	0.9974	C	G	-0.00653	0.004105	0.11185839	1736	0.469
0.4828	0.995	A	G	-0.00616	0.00406	0.1296977	1736	0.4828
0.7455	0.9996	C	T	0.007147	0.004684	0.12732665	1736	0.2545
0.5315	0.9994	A	G	0.00647	0.004105	0.11518703	1736	0.4685
0.5957	0.9954	A	T	-0.00671	0.004155	0.10672656	1736	0.4043
0.5302	0.9833	A	G	0.006446	0.004118	0.1177556	1736	0.4698
0.6026	0.9949	C	T	-0.00425	0.004146	0.30546344	1736	0.3974
0.4029	0.9923	A	G	0.006771	0.004169	0.10456591	1736	0.4029
0.5957	0.9946	A	G	-0.00669	0.004157	0.10767824	1736	0.4043
0.5999	0.9993	A	G	-0.0076	0.004158	0.06785587	1736	0.4001
0.5957	0.9953	C	G	-0.0067	0.004155	0.10688339	1736	0.4043

0.9647	0.9462	C	G	0.01668	0.01195	0.16286949	1736	0.0353
0.5958	0.9954	A	G	-0.00673	0.004155	0.10573157	1736	0.4042
0.4037	0.8171	C	G	0.005112	0.005397	0.34377563	1736	0.4037
0.1023	0.9672	C	T	-0.00649	0.006674	0.33064496	1736	0.1023
0.3678	0.924	C	T	-0.00429	0.004218	0.30901759	1736	0.3678

Framingham Heart Study

af_framhs2	oevar_framhs2	coded_framhs2	noncoded_framhs2	beta_framhs2	se_framhs2	p_framhs2	n_framhs2	af_genoa
0.466703	0.997683	A	G	0.004905069	0.002858255	0.0861426	3549	0.4863
0.466639	0.998217	G	C	0.004897496	0.002857536	0.08654938	3549	0.4873
0.455599	0.920944	G	C	-0.003855664	0.002989724	0.19717613	3549	0.5697
0.455712	0.920072	G	A	-0.003835981	0.002990884	0.19964704	3549	0.5698
0.449397	0.958738	T	C	0.004396344	0.002918468	0.13196766	3549	0.5107
0.4495	0.958574	A	G	0.004404341	0.00291861	0.13128484	3549	0.5105
0.449449	0.95865	A	G	0.004400584	0.002918543	0.13160533	3549	0.5107
0.449534	0.958527	A	T	0.004407311	0.002918645	0.13102973	3549	0.5104
0.449656	0.958372	T	C	0.004415748	0.00291875	0.13030751	3549	0.5102
0.448106	0.95202	T	C	0.004483664	0.002929153	0.12584272	3549	0.5121
0.448141	0.952006	A	C	0.004486314	0.002929145	0.12561813	3549	0.512
0.411342	0.998398	C	T	0.005979733	0.0029383	0.04184087	3549	0.5532
0.377222	0.993649	A	G	-0.002344397	0.002978229	0.43117752	3549	0.3681
0.449708	0.958314	A	G	0.004419984	0.002918792	0.12994473	3549	0.5101
0.449828	0.958199	A	G	0.00442874	0.002918843	0.12919274	3549	0.51
0.44988	0.958149	G	C	0.004432532	0.002918868	0.12886845	3549	0.5099
0.4535	0.977938	A	C	0.004029132	0.002895102	0.16401125	3549	0.5067
0.453431	0.978199	T	A	0.00402496	0.002894803	0.16440466	3549	0.5069
0.450361	0.957942	T	C	0.004465039	0.00291874	0.12607042	3549	0.5091
0.450241	0.957956	T	C	0.004456452	0.002918827	0.12681181	3549	0.5093
0.46108	0.981933	A	G	0.004518751	0.002886832	0.11751317	3549	0.4986
0.470976	0.991977	C	T	0.003865274	0.002866203	0.17747534	3549	0.4829
0.471348	0.988808	G	T	0.003902746	0.002870874	0.17401106	3549	0.4825
0.415136	0.999605	G	A	0.005870188	0.002929993	0.0451256	3549	0.5497
0.450191	0.069315	A	G	-0.012246108	0.010964361	0.26403654	3549	0.5258
0.333296	0.963527	C	T	0.010622616	0.003105571	0.00062505	3549	0.6339
0.35709	0.938903	A	G	-0.003034314	0.003097915	0.32734796	3549	0.3565
0.530102	1.00501	A	T	0.003550696	0.002892502	0.21961519	3549	0.4469
0.530113	1.005015	C	G	0.003551996	0.002892446	0.21943748	3549	0.447
0.466229	0.970623	T	A	0.003828443	0.002894915	0.18601111	3549	0.5014
0.388888	0.998091	C	T	0.00690759	0.002963121	0.01974348	3549	0.5654
0.389781	0.992495	G	T	0.006906015	0.002971304	0.02011277	3549	0.5642
0.39697	1.003092	G	A	0.00706545	0.002953205	0.01673547	3549	0.5577
0.436893	0.972533	G	C	0.006670763	0.002967709	0.0245901	3549	0.5382
0.387628	0.999586	T	A	0.006645419	0.002960919	0.02480816	3549	0.5659
0.398066	0.997607	C	T	0.006966644	0.00295976	0.01858326	3549	0.5575
0.387718	0.999754	A	G	0.006686441	0.002960233	0.02389862	3549	0.566
0.533653	1.004195	T	A	0.003438015	0.002893469	0.2347553	3549	0.4432
0.522836	0.994388	T	C	-0.003628264	0.002866581	0.20561648	3549	0.4898
0.392538	0.963575	G	T	0.00689074	0.003025975	0.02277449	3549	0.5674
0.468581	0.994921	A	C	0.003374079	0.002866863	0.23922597	3549	0.5001
0.528497	0.977801	G	A	-0.003882686	0.002887983	0.17880996	3549	0.4895
0.534252	0.999805	G	T	0.003515687	0.002901158	0.22558053	3549	0.4432
0.465827	1.000331	G	A	-0.003505773	0.002900221	0.22674154	3549	0.4432
0.471174	0.985317	C	T	0.003863648	0.002876423	0.17920309	3549	0.4865
0.464681	0.996104	T	G	-0.003615923	0.002906528	0.21347392	3549	0.4431
0.477462	0.995308	G	C	0.00390356	0.002864428	0.17295464	3549	0.491
0.477488	0.995764	C	A	0.003902902	0.002863872	0.17294367	3549	0.491
0.396824	0.984211	G	A	0.006277096	0.002979029	0.03510938	3549	0.5632
0.381607	0.987415	G	C	-0.002507419	0.002980842	0.40024826	3549	0.3724

0.522677	0.993528	T	C	-0.003904727	0.002866637	0.1731574	3549	0.4909
0.477272	0.993152	T	G	0.003903843	0.002867085	0.17332199	3549	0.4908
0.522461	0.996781	G	A	-0.003901233	0.002862587	0.17293461	3549	0.491
0.522339	0.996895	G	A	-0.003888573	0.00286269	0.17434844	3549	0.491
0.366731	0.978936	A	T	-0.002900952	0.003021388	0.33698541	3549	0.3682
0.392023	0.959906	C	T	0.006419964	0.003029127	0.03405578	3549	0.5672
0.402068	0.989816	C	T	0.006120332	0.002965435	0.0390281	3549	0.5569
0.36449	0.9292	T	C	-0.011029536	0.003062245	0.00031604	3549	0.3423
0.123755	0.863486	G	T	-0.003003164	0.004597228	0.51359159	3549	0.8529
0.405376	0.984303	G	A	0.007507584	0.002966938	0.01139278	3549	0.5534
0.123838	0.861833	C	A	-0.003014529	0.004600135	0.51226612	3549	0.8518
0.509436	0.983425	G	A	-0.00374517	0.002876468	0.1929154	3549	0.4876
0.509488	0.983702	T	A	-0.003741171	0.002876137	0.1933398	3549	0.4877
0.507096	0.997761	A	G	0.004436798	0.002882515	0.12375292	3549	0.4626
0.400548	1.002592	C	T	0.005783124	0.002945232	0.04958177	3549	0.5567
0.176637	0.997971	A	G	-0.004928936	0.003771491	0.19124899	3549	0.8327
0.364267	0.93116	C	T	-0.010928015	0.003058832	0.00035344	3549	0.3423
0.506264	1.001894	T	A	0.003968222	0.002874868	0.1674898	3549	0.4595
0.124512	0.850915	C	A	-0.003096821	0.004617825	0.50246056	3549	0.851
0.509521	0.983883	G	C	-0.003739164	0.002875919	0.19354499	3549	0.488
0.50953	0.985255	C	G	-0.003707415	0.002874341	0.19710919	3549	0.489
0.401086	1.004736	C	T	0.005723606	0.00294179	0.05170034	3549	0.5553
0.362017	0.921125	A	G	-0.010980243	0.003082647	0.00036811	3549	0.3423
0.135212	0.780571	G	A	-0.00338876	0.004652765	0.46641004	3549	0.841
0.400821	1.003897	G	C	0.00574951	0.002943177	0.05075987	3549	0.5558
0.418933	1.013299	G	A	-0.007354205	0.002870295	0.01040182	3549	0.4057
0.424478	1.015437	G	A	-0.007615966	0.002857745	0.00769824	3549	0.4022
0.50123	0.991152	G	A	0.004433182	0.002886716	0.12460745	3549	0.4632
0.497866	0.966175	G	A	0.004490948	0.002929679	0.12529689	3549	0.4725
0.502147	0.988425	T	C	0.004281555	0.002891187	0.13863358	3549	0.4623
0.506432	1.002265	G	C	0.004058601	0.002874317	0.15794332	3549	0.4598
0.485519	0.949425	G	A	0.004534653	0.002947117	0.12388379	3549	0.4687
0.245306	0.839369	T	C	-0.007918995	0.003620226	0.02871107	3549	0.7656
0.506479	1.002479	A	G	0.004082759	0.002873837	0.15541427	3549	0.46
0.365328	0.985371	T	A	0.006877507	0.003026207	0.02304735	3549	0.4071
0.498552	0.998731	A	G	0.004147073	0.002873766	0.14899804	3549	0.463
0.387225	0.910888	T	C	-0.00672056	0.003085659	0.02940614	3549	0.6044
0.364518	0.982477	A	G	0.00691069	0.003032621	0.02267988	3549	0.5937
0.36538	0.985834	G	A	0.006895828	0.003025464	0.02265143	3549	0.4067
0.365277	0.984626	G	A	0.006858445	0.003027312	0.02348061	3549	0.4069
0.365336	0.985507	G	C	0.0068811	0.003025989	0.02296596	3549	0.407
0.036004	0.773943	G	C	-0.025800581	0.008900193	0.00374499	3549	0.9694
0.365291	0.984796	G	A	0.006862013	0.003027082	0.02339792	3549	0.407
0.41123	0.985633	C	G	0.005723602	0.002948515	0.05223628	3549	0.6042
0.110229	0.888246	C	T	-0.000303995	0.004782559	0.94931804	3549	0.8794
0.363676	1.02407	C	T	-0.009152632	0.002933134	0.00180589	3549	0.3252

GENOA

af_genoa	oevar_genoa	coded_genoa	noncoded_genoa	beta_genoa	se_genoa	p_genoa	n_genoa	maf_genoa
0.4863	0.981759369	G	A	-0.009542862	0.00620261	0.1244874	991	0.4863
0.4873	0.988681355	C	G	-0.009479123	0.00618203	0.12575985	991	0.4873
0.5697	0.962113333	C	G	0.014537652	0.00622841	0.01994525	991	0.4303
0.5698	0.96250553	A	G	0.014560348	0.006226	0.01970631	991	0.4302
0.5107	0.99662403	C	T	-0.006133881	0.00613526	0.31785224	991	0.4893
0.5105	0.996349189	G	A	-0.006180971	0.00613673	0.31427044	991	0.4895
0.5107	0.996534959	G	A	-0.006149726	0.00613575	0.31664349	991	0.4893
0.5104	0.996278236	T	A	-0.006196835	0.00613717	0.31306644	991	0.4896
0.5102	0.996129369	C	T	-0.006232963	0.00613809	0.31032868	991	0.4898
0.5121	0.987623696	C	T	-0.006291712	0.00616262	0.30772145	991	0.4879
0.512	0.987654045	C	A	-0.006309808	0.00616283	0.30635018	991	0.488
0.5532	0.974664823	T	C	-0.006851864	0.00609122	0.26112534	991	0.4468
0.3681	0.950141706	A	G	-0.007129499	0.0063466	0.26176862	991	0.3681
0.5101	0.99609431	G	A	-0.006254734	0.00613846	0.30867201	991	0.4899
0.51	0.996043672	G	A	-0.006272175	0.00613888	0.30735913	991	0.49
0.5099	0.995997014	C	G	-0.006302068	0.00613946	0.30510729	991	0.4901
0.5067	1.0178619	C	A	-0.006657216	0.00607013	0.27323822	991	0.4933
0.5069	1.018193158	A	T	-0.006628104	0.00606885	0.27523769	991	0.4931
0.5091	0.996142806	C	T	-0.006452383	0.00614125	0.2938682	991	0.4909
0.5093	0.99602494	C	T	-0.006411272	0.00614095	0.29692854	991	0.4907
0.4986	1.030161383	G	A	-0.007929684	0.00605621	0.19095541	991	0.4986
0.4829	1.049148243	T	C	-0.008922056	0.00600216	0.13771847	991	0.4829
0.4825	1.04063442	T	G	-0.008996756	0.00602774	0.13611778	991	0.4825
0.5497	0.978463748	A	G	-0.006064883	0.00610303	0.32077589	991	0.4503
0.5258	0.583574908	G	A	0.008970029	0.00786643	0.25465386	991	0.4742
0.6339	0.98487417	T	C	-0.007876609	0.00636312	0.21629126	991	0.3661
0.3565	0.876746955	A	G	-0.006389075	0.00664383	0.33663935	991	0.3565
0.4469	0.963379076	T	A	-0.010233476	0.00607905	0.09285692	991	0.4469
0.447	0.963172909	G	C	-0.01024111	0.00607938	0.09263186	991	0.447
0.5014	1.020953815	A	T	-0.005599969	0.00606525	0.35625663	991	0.4986
0.5654	0.976530355	T	C	-0.007462287	0.00614478	0.2251048	991	0.4346
0.5642	0.979150639	T	G	-0.0072361	0.00613899	0.2390153	991	0.4358
0.5577	0.983301954	A	G	-0.00726093	0.00610515	0.23482281	991	0.4423
0.5382	0.959663973	C	G	-0.005742609	0.00614494	0.35043668	991	0.4618
0.5659	0.973909284	A	T	-0.007096312	0.00615266	0.24925123	991	0.4341
0.5575	0.982561719	T	C	-0.007232353	0.0061062	0.23674809	991	0.4425
0.566	0.973584247	G	A	-0.007132041	0.00615352	0.24694454	991	0.434
0.4432	0.963633804	A	T	-0.009275117	0.00610872	0.12949502	991	0.4432
0.4898	1.045988631	T	C	-0.007627454	0.0060044	0.20450305	991	0.4898
0.5674	0.950094544	T	G	-0.00807885	0.00620773	0.19365247	991	0.4326
0.5001	1.047122821	C	A	-0.005604953	0.00598261	0.34922829	991	0.4999
0.4895	1.040880369	G	A	-0.007725695	0.00601887	0.19982172	991	0.4895
0.4432	0.962886013	T	G	-0.009267582	0.00611122	0.12996289	991	0.4432
0.4432	0.963125879	G	A	-0.009270137	0.00611045	0.12980895	991	0.4432

0.4865	1.01637299	T	C	-0.008488613	0.00609124	0.1639999	991	0.4865
0.4431	0.960419465	T	G	-0.009237497	0.00611931	0.13172025	991	0.4431
0.491	1.056149526	C	G	-0.00731755	0.00598023	0.22161068	991	0.491
0.491	1.056392444	A	C	-0.007356491	0.00597902	0.21907187	991	0.491
0.5632	0.977157879	A	G	-0.007179545	0.00614014	0.24278992	991	0.4368
0.3724	0.961815082	G	C	-0.006855629	0.00628504	0.27583845	991	0.3724
0.4909	1.055566173	T	C	-0.007185697	0.00598329	0.23027618	991	0.4909
0.4908	1.055503224	G	T	-0.007161357	0.0059837	0.23188888	991	0.4908
0.491	1.056796941	G	A	-0.007416123	0.00597699	0.21520848	991	0.491
0.491	1.056796941	G	A	-0.007416123	0.00597699	0.21520848	991	0.491
0.3682	0.950345456	A	T	-0.006826133	0.00634301	0.28231796	991	0.3682
0.5672	0.949697511	T	C	-0.008055007	0.00620784	0.19497665	991	0.4328
0.5569	0.981589653	T	C	-0.007173985	0.00610623	0.24055034	991	0.4431
0.3423	0.95902294	T	C	-0.014702838	0.00646229	0.02327397	991	0.3423
0.8529	0.971694588	T	G	0.014964336	0.00886	0.09178257	991	0.1471
0.5534	0.957846425	A	G	-0.006094871	0.00619895	0.32593029	991	0.4466
0.8518	0.974176825	A	C	0.015172137	0.00881689	0.08584066	991	0.1482
0.4876	1.027478264	G	A	-0.008677053	0.00604169	0.15150659	991	0.4876
0.4877	1.027900793	T	A	-0.008648809	0.00604012	0.15273434	991	0.4877
0.4626	0.959211536	G	A	-0.010369677	0.00615149	0.0924094	991	0.4626
0.5567	0.981336985	T	C	-0.008229459	0.00610022	0.17787001	991	0.4433
0.8327	0.989721001	G	A	0.010040433	0.00807887	0.21446267	991	0.1673
0.3423	0.959213314	C	T	-0.014682644	0.00646177	0.02345212	991	0.3423
0.4595	0.961611957	A	T	-0.010451854	0.00611791	0.08811778	991	0.4595
0.851	0.977679412	A	C	0.015294845	0.00877848	0.08200498	991	0.149
0.488	1.028801707	G	C	-0.008588507	0.0060368	0.15538445	991	0.488
0.489	1.026798416	C	G	-0.009039627	0.00604214	0.13519386	991	0.489
0.5553	0.985024179	T	C	-0.008349062	0.00608664	0.17070702	991	0.4447
0.3423	0.959121894	A	G	-0.014745881	0.00646193	0.02286718	991	0.3423
0.841	0.918961991	A	G	0.015350795	0.0088113	0.08202943	991	0.159
0.5558	0.98508204	C	G	-0.008311062	0.00608691	0.17267755	991	0.4442
0.4057	1.035723921	G	A	-0.009042601	0.00615409	0.14229765	991	0.4057
0.4022	1.023301062	G	A	-0.008499748	0.00618823	0.17013814	991	0.4022
0.4632	0.956379819	A	G	-0.012002819	0.00613147	0.05077766	991	0.4632
0.4725	0.925190212	A	G	-0.011302016	0.00624673	0.07094726	991	0.4725
0.4623	0.955114774	C	T	-0.010340861	0.00615975	0.09375481	991	0.4623
0.4598	0.96043873	C	G	-0.010515078	0.00612	0.08632377	991	0.4598
0.4687	0.944465156	A	G	-0.014489293	0.00618557	0.01950885	991	0.4687
0.7656	0.868435534	C	T	0.010311953	0.00764507	0.17793658	991	0.2344
0.46	0.960358743	G	A	-0.010531966	0.00611985	0.08581384	991	0.46
0.4071	1.014870036	T	A	0.009577611	0.00615081	0.12000671	991	0.4071
0.463	0.958532439	G	A	-0.012013566	0.00612424	0.05030086	991	0.463
0.6044	1.031328457	C	T	0.010646792	0.00610245	0.08159215	991	0.3956
0.5937	1.01121449	G	A	-0.009466142	0.00615513	0.12463345	991	0.4063
0.4067	1.014959934	G	A	0.009508945	0.00615151	0.12272189	991	0.4067
0.4069	1.004238981	G	A	0.00968597	0.00618678	0.11801104	991	0.4069
0.407	1.014828139	G	C	0.009567811	0.00615164	0.12043516	991	0.407

0.9694	0.830894535	C	G	0.006508432	0.02048663	0.75083768	991	0.0306
0.407	1.011229147	G	A	0.009617301	0.00616318	0.11922163	991	0.407
0.6042	0.97072276	G	C	-0.007719134	0.00633031	0.22321013	991	0.3958
0.8794	0.946478265	T	C	0.014810571	0.0099366	0.13665607	991	0.1206
0.3252	0.896550839	C	T	-0.017106249	0.00673357	0.01134115	991	0.3252

Gutenberg Heart Study

af_guten	oevar_guten	coded_guten	noncoded_guten	beta_guten	se_guten	p_guten	n_guten	maf_guten
0.483345	0.949711	G	A	-0.00329573	0.00299794	0.271174	2762	0.483345
0.51648	0.949714	G	C	0.00329188	0.00299809	0.272534	2761	0.48352
0.430841	0.976687	G	C	-0.00459936	0.00298117	0.123381	3022	0.430841
0.430818	0.976581	G	A	-0.00458917	0.00298141	0.123783	3021	0.430818
0.500166	0.982863	T	C	0.00267697	0.00295265	0.364234	3018	0.499834
0.499834	0.983022	G	A	-0.00268304	0.00295254	0.363996	3018	0.499834
0.499834	0.982964	G	A	-0.0026777	0.00295258	0.364758	3018	0.499834
0.499834	0.983015	T	A	-0.00268446	0.00295253	0.363792	3018	0.499834
0.500166	0.982901	T	C	0.00268318	0.00295252	0.363055	3018	0.499834
0.499336	0.980314	T	C	0.00264877	0.00295529	0.369483	3013	0.499336
0.500664	0.980435	C	A	-0.00265082	0.00295527	0.370689	3013	0.499336
0.557395	0.998143	T	C	-0.00802581	0.00296013	0.00669655	3110	0.442605
0.636231	0.99057	G	A	0.0058553	0.00302364	0.0526642	3083	0.363769
0.499834	0.983019	G	A	-0.00268318	0.00295252	0.363922	3019	0.499834
0.499834	0.983036	G	A	-0.00268261	0.00295252	0.363839	3019	0.499834
0.500166	0.982901	G	C	0.00268261	0.00295252	0.363278	3019	0.499834
0.498534	0.992505	C	A	-0.00186827	0.00293835	0.524794	3069	0.498534
0.501466	0.992506	T	A	0.00186905	0.00293835	0.524982	3069	0.498534
0.500331	0.983102	T	C	0.00268665	0.00295221	0.362741	3020	0.499669
0.500166	0.982999	T	C	0.00268471	0.00295234	0.362747	3019	0.499834
0.496622	1	G	A	-0.002708	0.00293237	0.355755	3108	0.496622
0.478059	1	T	C	-0.00343934	0.00291202	0.237569	3122	0.478059
0.478106	0.97612	T	G	-0.002645	0.00294918	0.369473	2946	0.478106
0.44811	1	G	A	0.00699285	0.00296009	0.018158	3122	0.44811
0.512484	1	G	A	0.0110502	0.00302661	0.00026121	3084	0.487516
0.660289	0.996976	T	C	-0.00467105	0.00304432	0.124854	3110	0.339711
0.653435	0.926743	G	A	0.00432443	0.00317245	0.173581	2809	0.346565
0.450322	0.998302	T	A	-0.00769645	0.00292851	0.00858382	3110	0.450322
0.450322	0.998261	G	C	-0.00769546	0.00292861	0.00859451	3110	0.450322
0.509622	0.988908	T	A	0.00170317	0.00293707	0.561746	3066	0.490378
0.582745	1	T	C	-0.00688951	0.00298734	0.0210974	3118	0.417255
0.582663	0.997193	T	G	-0.00670866	0.00298332	0.0246039	3109	0.417337
0.42729	1	G	A	0.00658816	0.00296662	0.0263669	3122	0.42729
0.45664	0.968127	G	C	0.00697903	0.00298637	0.0197514	2952	0.45664
0.41664	0.99853	T	A	0.00680192	0.0029843	0.022653	3113	0.41664
0.572283	0.995896	T	C	-0.00672701	0.00296861	0.0234498	3092	0.427717
0.582906	0.999128	G	A	-0.00674789	0.00298322	0.0237006	3118	0.417094
0.554397	1	T	A	0.00671661	0.00294305	0.022478	3116	0.445603
0.480942	1	T	C	-0.00171138	0.00292488	0.558473	3122	0.480942
0.587599	0.962627	T	G	-0.00669981	0.00302306	0.026794	2911	0.412401
0.490391	1	C	A	-0.000635501	0.00291898	0.827653	3122	0.490391
0.482532	0.998653	G	A	-0.00160637	0.00292667	0.583174	3120	0.482532
0.446371	0.99657	T	G	-0.00665433	0.00293954	0.0235341	3114	0.446371
0.446406	0.997359	G	A	-0.00664759	0.00293876	0.0236497	3116	0.446406

0.477642	0.978838	T	C	-0.00196066	0.0029538	0.506125	2952	0.477642
0.445458	0.989954	T	G	-0.00663538	0.00294615	0.0241749	3071	0.445458
0.517485	0.99943	G	C	0.00131875	0.00292429	0.652015	3117	0.482515
0.517485	0.99943	C	A	0.00131875	0.00292429	0.652015	3117	0.482515
0.418081	0.990776	G	A	0.00677482	0.00298567	0.0232138	3064	0.418081
0.366753	0.989062	G	C	-0.0046312	0.00303179	0.126668	3077	0.366753
0.482515	0.99943	T	C	-0.00131875	0.00292429	0.652015	3117	0.482515
0.517485	0.99943	T	G	0.00131875	0.00292429	0.652015	3117	0.482515
0.481897	1	G	A	-0.00120561	0.00292304	0.680011	3121	0.481897
0.482223	1	G	A	-0.00118099	0.00292624	0.686518	3122	0.482223
0.644485	0.968986	T	A	0.00492175	0.00308321	0.11128	2983	0.355515
0.588458	0.962169	T	C	-0.0068063	0.00302531	0.0244193	2911	0.411542
0.57294	0.991634	T	C	-0.00684931	0.00297136	0.0211454	3071	0.42706
0.361958	0.998717	T	C	-0.00671244	0.0030382	0.0271905	3115	0.361958
0.885336	0.95836	T	G	0.00877027	0.00457112	0.0554806	2987	0.114664
0.423968	0.964599	G	A	0.00556912	0.00301998	0.0655882	2979	0.423968
0.117667	0.96343	C	A	-0.00866668	0.00454746	0.0566928	3000	0.117667
0.468991	0.996595	G	A	0.00038263	0.00293302	0.895683	3112	0.468991
0.468991	0.996595	T	A	0.00038263	0.00293302	0.895683	3112	0.468991
0.477523	0.994981	G	A	-0.00605932	0.00293004	0.0386459	3092	0.477523
0.577224	0.993559	T	C	-0.00742349	0.00296329	0.0122567	3069	0.422776
0.827936	0.982831	G	A	0.00389456	0.00390793	0.318038	3057	0.172064
0.637981	0.999475	T	C	0.0062398	0.00303229	0.0395878	3120	0.362019
0.522742	1	T	A	0.00657705	0.00292853	0.0247134	3122	0.477258
0.12	0.972212	C	A	-0.00854873	0.00451141	0.0580966	3025	0.12
0.468991	0.996595	G	C	0.000382981	0.00293303	0.895682	3112	0.468991
0.53075	1	G	C	-0.000382205	0.0029264	0.896087	3122	0.46925
0.573335	0.996035	T	C	-0.00746659	0.00295836	0.0115838	3109	0.426665
0.639704	1	G	A	0.00604796	0.00304655	0.0471239	3103	0.360296
0.12205	0.905476	G	A	-0.00743565	0.00451165	0.0990507	2839	0.12205
0.42623	0.996824	G	C	0.00744467	0.00295831	0.0118476	3111	0.42623
0.426603	0.999603	G	A	-0.00631136	0.00298904	0.0346744	3120	0.426603
0.429311	0.99587	G	A	-0.00616046	0.0029905	0.0393977	3091	0.429311
0.517179	0.992104	G	A	0.00670428	0.00293252	0.0222467	3056	0.482821
0.508408	0.963538	G	A	0.00593915	0.00297256	0.0457891	2914	0.491592
0.52082	0.992095	T	C	0.00610614	0.00293599	0.0375036	3074	0.47918
0.523771	0.998566	G	C	0.00637252	0.0029296	0.029584	3113	0.476229
0.516522	1	G	A	0.00707238	0.0029359	0.0159993	3117	0.483478
0.226308	0.909396	T	C	-0.00664467	0.00363747	0.0679219	2733	0.226308
0.476076	0.999097	G	A	-0.00631491	0.00292907	0.0310881	3114	0.476076
0.408535	1	T	A	0.00521936	0.00297997	0.0798627	3105	0.408535
0.483643	1	G	A	-0.0068524	0.00292742	0.0192442	3118	0.483643
0.38291	1	T	C	-0.00886497	0.00303062	0.00344307	3113	0.38291
0.593992	0.991206	G	A	-0.005417	0.00297144	0.0681959	3096	0.406008
0.406611	0.998796	G	A	0.00537775	0.0029668	0.0698445	3116	0.406611
0.406581	0.998757	G	A	0.00536501	0.00296695	0.070522	3115	0.406581
0.406611	0.998878	G	C	0.00537177	0.00296672	0.0701244	3116	0.406611

0.0106275	0.77367	G	C	-0.0110879	0.0104711	0.289678	2964	0.0106275
0.406581	0.998855	G	A	0.0053661	0.00296679	0.0704378	3115	0.406581
0.554484	1	G	C	-0.006276	0.00296811	0.0344744	3111	0.445516
0.9040522	0.949436	T	C	0.0127457	0.00489628	0.00930972	2986	0.0959478
0.643177	1	T	C	0.00536654	0.00301971	0.0755404	3122	0.356823

Health ABC

af_habc	oevar_habc	coded_habc	noncoded_habc	beta_habc	se_habc	p_habc	n_habc	maf_habc
0.4988	0.9767	A	G	0.015465851	0.00549687	0.00496083	1564	0.4988
0.5012	0.9768	C	G	-0.01547994	0.00549609	0.00491559	1564	0.4988
0.5601	0.9991	C	G	0.009774864	0.00548397	0.07487148	1564	0.4399
0.5602	0.9994	A	G	0.009818172	0.00548369	0.07357843	1564	0.4398
0.533	0.9962	C	T	-0.01697609	0.00544527	0.00185669	1564	0.467
0.466	0.9996	A	G	0.016939701	0.00543384	0.00185758	1564	0.466
0.4662	0.9988	A	G	0.016948715	0.00543632	0.00185609	1564	0.4662
0.4695	0.9991	A	T	0.016604412	0.00546383	0.00241333	1564	0.4695
0.5303	0.9995	C	T	-0.01658508	0.00546355	0.00244043	1564	0.4697
0.5357	0.9746	C	T	-0.01685806	0.00553179	0.00234639	1564	0.4643
0.4644	0.9743	A	C	0.01686088	0.00553258	0.00234579	1564	0.4644
0.4264	0.9976	C	T	0.006355011	0.00558542	0.25538495	1564	0.4264
0.3721	0.9989	A	G	-0.0058316	0.00571457	0.30765963	1564	0.3721
0.4698	0.999	A	G	0.016590823	0.00546499	0.00243841	1564	0.4698
0.47	0.9981	A	G	0.016600372	0.00546722	0.00243431	1564	0.47
0.53	0.9978	C	G	-0.01660268	0.00546778	0.0024334	1564	0.47
0.4785	0.9362	A	C	0.018451751	0.0056218	0.00105285	1564	0.4785
0.5215	0.9349	A	T	-0.01846562	0.0056257	0.00105215	1564	0.4785
0.5282	0.9908	C	T	-0.0166787	0.00548646	0.00240536	1564	0.4718
0.5284	0.9917	C	T	-0.01667059	0.00548407	0.00240657	1564	0.4716
0.4912	0.9786	A	G	0.017779993	0.00552786	0.00132454	1564	0.4912
0.504	0.9986	C	T	0.016088001	0.00544253	0.0031636	1564	0.496
0.5036	0.9989	G	T	0.016123886	0.00543717	0.00306787	1564	0.4964
0.5702	0.9971	A	G	-0.00621093	0.005564	0.2644784	1564	0.4298
0.4816	0.9962	A	G	-0.00736396	0.00546745	0.17821637	1564	0.4816
0.3328	0.9994	C	T	-0.00025898	0.00574764	0.96406663	1564	0.3328
0.3644	0.9981	A	G	-0.00338128	0.00568453	0.55204841	1564	0.3644
0.5298	0.9985	A	T	0.001368799	0.00554038	0.80489519	1564	0.4702
0.5298	0.9986	C	G	0.001379691	0.00553983	0.80335544	1564	0.4702
0.5225	0.9787	A	T	-0.01719536	0.00552333	0.00188417	1564	0.4775
0.4057	0.9995	C	T	0.004349059	0.00562649	0.43966253	1564	0.4057
0.4061	0.9994	G	T	0.004379425	0.00561818	0.4357976	1564	0.4061
0.5873	0.999	A	G	-0.00465772	0.00561084	0.40659248	1564	0.4127
0.5488	0.9828	C	G	-0.00455818	0.00554506	0.41118787	1564	0.4512
0.5954	0.9992	A	T	-0.00408965	0.00562809	0.46755017	1564	0.4046
0.4131	0.9999	C	T	0.00469813	0.00559852	0.40150048	1564	0.4131
0.4047	0.9995	A	G	0.004127227	0.00562752	0.46342404	1564	0.4047
0.4667	0.9987	A	T	-0.00126797	0.00551297	0.81812329	1564	0.4667
0.5021	0.9997	C	T	0.018319584	0.00550432	0.00089438	1564	0.4979
0.4131	0.9994	G	T	0.004658327	0.00560054	0.4056693	1564	0.4131
0.4881	0.8547	A	C	0.019448709	0.00588251	0.00096722	1564	0.4881
0.5021	0.9997	A	G	0.018319689	0.0055043	0.00089429	1564	0.4979
0.5334	0.9991	G	T	0.001271287	0.00551064	0.81758051	1564	0.4666
0.5332	0.9995	A	G	0.001237227	0.00551092	0.82239422	1564	0.4668
0.5045	0.9994	C	T	0.016048518	0.00545225	0.00329355	1564	0.4955

0.5341	0.9988	G	T	0.001343881	0.00550653	0.80722238	1564	0.4659
0.5016	0.9968	C	G	-0.0177827	0.00548721	0.00121728	1564	0.4984
0.5014	0.9977	A	C	-0.01776596	0.00548527	0.00122539	1564	0.4986
0.5942	0.9993	A	G	-0.00406598	0.00562273	0.46970684	1564	0.4058
0.6243	0.9985	C	G	0.005692442	0.00566271	0.31493181	1564	0.3757
0.4981	0.9954	C	T	0.017805304	0.00549004	0.00120693	1564	0.4981
0.502	0.9951	G	T	-0.01781139	0.00549073	0.001204	1564	0.498
0.4987	0.998	A	G	0.017760653	0.00548457	0.00122775	1564	0.4987
0.4988	0.9987	A	G	0.017748423	0.00548305	0.00123345	1564	0.4988
0.3689	0.9768	A	T	-0.00418291	0.00574109	0.46636088	1564	0.3689
0.4125	0.9971	C	T	0.004447637	0.00561187	0.42816589	1564	0.4125
0.4125	0.9982	C	T	0.004296833	0.00560619	0.44352783	1564	0.4125
0.6462	0.9971	C	T	-0.00227996	0.00568769	0.68857953	1564	0.3538
0.1064	1	G	T	-0.00064761	0.00904172	0.94291019	1564	0.1064
0.5806	0.9988	A	G	-0.00217403	0.00554394	0.69500396	1564	0.4194
0.8936	1	A	C	0.000647606	0.00904172	0.94291019	1564	0.1064
0.4951	0.8731	A	G	0.018571337	0.0057988	0.00138933	1564	0.4951
0.4953	0.8708	A	T	0.018584613	0.00580626	0.00139812	1564	0.4953
0.5123	0.9985	A	G	-0.00045822	0.00551594	0.93380552	1564	0.4877
0.4206	0.9979	C	T	0.006419806	0.00559373	0.2512772	1564	0.4206
0.1724	0.9991	A	G	-0.01209966	0.00735067	0.0999524	1564	0.1724
0.3548	1	C	T	0.002045112	0.00567544	0.7186384	1564	0.3548
0.4921	0.9981	A	T	0.000943139	0.00551726	0.86429048	1564	0.4921
0.8936	1	A	C	0.000647606	0.00904172	0.94291019	1564	0.1064
0.4959	0.8624	C	G	0.018617195	0.00583402	0.00144531	1564	0.4959
0.5028	0.8452	C	G	-0.01868205	0.00589226	0.00155078	1564	0.4972
0.421	0.9952	C	T	0.006395927	0.00560141	0.25369461	1564	0.421
0.354	0.9949	A	G	0.002544244	0.00568641	0.65462992	1564	0.354
0.8802	0.8479	A	G	-0.00046643	0.00926907	0.95987256	1564	0.1198
0.5792	0.9969	C	G	-0.00640714	0.00559663	0.25245859	1564	0.4208
0.5778	0.9992	A	G	0.006406946	0.0055641	0.24971125	1564	0.4222
0.5772	0.9996	A	G	0.006481399	0.00556258	0.24412459	1564	0.4228
0.4896	0.9996	A	G	-0.00112176	0.00550232	0.83848215	1564	0.4896
0.4877	0.999	A	G	0.000522401	0.00551393	0.92453187	1564	0.4877
0.4885	0.9971	C	T	0.000828704	0.00551375	0.88054916	1564	0.4885
0.4919	0.9991	C	G	0.00086663	0.00551649	0.87518783	1564	0.4919
0.5016	0.9989	A	G	-0.00311985	0.00548572	0.56962624	1564	0.4984
0.7574	0.9994	C	T	0.008832601	0.0063791	0.16636786	1564	0.2426
0.5081	0.9994	A	G	-0.00084858	0.00551622	0.87776138	1564	0.4919
0.6065	0.9989	A	T	-0.00863443	0.00559895	0.12323954	1564	0.3935
0.5128	0.9913	A	G	0.000690864	0.00552981	0.90059148	1564	0.4872
0.601	1	C	T	0.006349893	0.00554872	0.25263816	1564	0.399
0.3931	0.9951	A	G	0.008522789	0.00561009	0.1289174	1564	0.3931
0.6065	0.9986	A	G	-0.00864219	0.00559982	0.12296125	1564	0.3935
0.6142	0.9999	A	G	-0.00783177	0.00561768	0.163477	1564	0.3858
0.6065	0.9988	C	G	-0.00863543	0.00559913	0.12320834	1564	0.3935
0.9708	0.9655	C	G	0.008154465	0.0162193	0.61520187	1564	0.0292

0.6068	0.9975	A	G	-0.0086085	0.00560383	0.12469691	1564	0.3932
0.3688	0.8623	C	G	0.006167875	0.00607977	0.31050556	1564	0.3688
0.0882	0.9543	C	T	-0.00093303	0.01003655	0.92594438	1564	0.0882
0.3622	0.9959	C	T	0.005511626	0.00567552	0.33163749	1564	0.3622

InCHIANTI

af_inch2	oevar_inch2	coded_inch2	noncoded_inch2	beta_inch2	se_inch2	p_inch2	n_inch2	maf_inch2
0.567	0.9546	G	A	-0.016	0.006	0.01698	1130	0.433
0.4328	0.9548	G	C	0.015	0.006	0.01712	1130	0.4328
0.4952	0.9929	G	C	-0.015	0.006	0.01778	1130	0.4952
0.4956	0.9928	G	A	-0.015	0.006	0.01782	1130	0.4956
0.4068	0.9887	T	C	0.018	0.006	0.005233	1130	0.4068
0.596	0.9989	G	A	-0.018	0.006	0.005359	1130	0.404
0.5949	0.9946	G	A	-0.018	0.006	0.005301	1130	0.4051
0.5944	0.9971	T	A	-0.018	0.006	0.005741	1130	0.4056
0.4062	0.9981	T	C	0.018	0.006	0.005941	1130	0.4062
0.4009	0.9745	T	C	0.017	0.006	0.007205	1130	0.4009
0.5988	0.9731	C	A	-0.017	0.006	0.007494	1130	0.4012
0.6446	0.9562	T	C	-0.015	0.007	0.02304	1130	0.3554
0.5701	0.9408	G	A	0.015	0.007	0.0265	1130	0.4299
0.5938	0.9976	G	A	-0.017	0.006	0.006015	1130	0.4062
0.5936	0.997	G	A	-0.017	0.006	0.006118	1130	0.4064
0.4065	0.996	G	C	0.017	0.006	0.006292	1130	0.4065
0.5826	0.8922	C	A	-0.019	0.007	0.005144	1130	0.4174
0.4174	0.8913	T	A	0.019	0.007	0.005175	1130	0.4174
0.4073	0.9921	T	C	0.017	0.006	0.007056	1130	0.4073
0.4072	0.9926	T	C	0.017	0.006	0.006941	1130	0.4072
0.5781	0.9684	G	A	-0.016	0.006	0.01276	1130	0.4219
0.5604	0.9961	T	C	-0.016	0.006	0.01413	1130	0.4396
0.5606	0.9965	T	G	-0.016	0.006	0.01367	1130	0.4394
0.3593	0.9532	G	A	0.015	0.007	0.02223	1130	0.3593
0.5316	0.7745	G	A	0.015	0.007	0.03414	1130	0.4684
0.7134	0.9995	T	C	-0.014	0.007	0.04043	1130	0.2866
0.5888	0.904	G	A	0.013	0.007	0.05321	1130	0.4112
0.5304	0.9481	T	A	-0.018	0.006	0.0056	1130	0.4696
0.5303	0.9482	G	C	-0.018	0.006	0.005605	1130	0.4697
0.4137	0.9759	T	A	0.017	0.006	0.009999	1130	0.4137
0.6614	0.9963	T	C	-0.013	0.007	0.04915	1130	0.3386
0.661	0.9977	T	G	-0.013	0.007	0.05017	1130	0.339
0.3459	0.9966	G	A	0.014	0.007	0.03221	1130	0.3459
0.375	0.9454	G	C	0.016	0.007	0.01423	1130	0.375
0.3379	0.9963	T	A	0.013	0.007	0.04918	1130	0.3379
0.6537	0.9985	T	C	-0.014	0.007	0.0329	1130	0.3463
0.6619	0.9969	G	A	-0.013	0.007	0.04872	1130	0.3381
0.471	0.9475	T	A	0.018	0.006	0.004408	1130	0.471
0.5693	0.9714	T	C	-0.015	0.006	0.0221	1130	0.4307
0.6537	0.9981	T	G	-0.014	0.007	0.03321	1130	0.3463
0.5702	0.8096	C	A	-0.018	0.007	0.008654	1130	0.4298
0.5693	0.9708	G	A	-0.015	0.006	0.02204	1130	0.4307
0.5293	0.9479	T	G	-0.018	0.006	0.004395	1130	0.4707
0.5292	0.9478	G	A	-0.018	0.006	0.004394	1130	0.4708

0.5602	0.9918	T	C	-0.015	0.006	0.01545	1130	0.4398
0.5312	0.9544	T	G	-0.018	0.006	0.004375	1130	0.4688
0.4295	0.9793	G	C	0.015	0.006	0.02252	1130	0.4295
0.4296	0.9797	C	A	0.015	0.006	0.02284	1130	0.4296
0.339	0.9968	G	A	0.013	0.007	0.05012	1130	0.339
0.4314	0.9452	G	C	-0.015	0.007	0.02052	1130	0.4314
0.5707	0.9786	T	C	-0.015	0.006	0.02169	1130	0.4293
0.4292	0.9784	T	G	0.015	0.006	0.02162	1130	0.4292
0.5703	0.9802	G	A	-0.014	0.006	0.02335	1130	0.4297
0.5702	0.9805	G	A	-0.014	0.006	0.02361	1130	0.4298
0.5771	0.9374	T	A	0.015	0.007	0.01842	1130	0.4229
0.6544	0.9951	T	C	-0.014	0.007	0.03085	1130	0.3456
0.6543	0.9955	T	C	-0.014	0.007	0.03047	1130	0.3457
0.3835	0.9986	T	C	-0.007	0.006	0.2986	1130	0.3835
0.8492	1	T	G	0.025	0.009	0.004779	1130	0.1508
0.3508	0.9995	G	A	0.012	0.007	0.06272	1130	0.3508
0.1508	1	C	A	-0.025	0.009	0.004778	1130	0.1508
0.5564	0.8414	G	A	-0.018	0.007	0.008343	1130	0.4436
0.5558	0.8355	T	A	-0.018	0.007	0.008414	1130	0.4442
0.5461	0.9844	G	A	-0.015	0.006	0.01446	1130	0.4539
0.6534	0.9745	T	C	-0.013	0.007	0.05715	1130	0.3466
0.7823	0.993	G	A	0.004	0.008	0.6397	1130	0.2177
0.6158	0.9994	T	C	0.007	0.006	0.3018	1130	0.3842
0.452	0.9824	T	A	0.016	0.006	0.01376	1130	0.452
0.1508	1	C	A	-0.025	0.009	0.004778	1130	0.1508
0.5553	0.8302	G	C	-0.018	0.007	0.008565	1130	0.4447
0.4483	0.8002	G	C	0.018	0.007	0.00973	1130	0.4483
0.6528	0.9732	T	C	-0.013	0.007	0.05805	1130	0.3472
0.6175	0.9953	G	A	0.007	0.006	0.2965	1130	0.3825
0.1598	0.9304	G	A	-0.026	0.009	0.003281	1130	0.1598
0.3469	0.9738	G	C	0.013	0.007	0.05762	1130	0.3469
0.4748	0.9999	G	A	-0.004	0.006	0.4906	1130	0.4748
0.4748	1	G	A	-0.004	0.006	0.491	1130	0.4748
0.4471	0.975	G	A	0.015	0.006	0.02333	1130	0.4471
0.4541	0.9847	G	A	0.015	0.006	0.01481	1130	0.4541
0.4517	0.9825	T	C	0.016	0.006	0.01347	1130	0.4517
0.4527	0.9839	G	C	0.016	0.006	0.01387	1130	0.4527
0.4368	0.9806	G	A	0.013	0.006	0.03867	1130	0.4368
0.2863	0.9961	T	C	-0.001	0.007	0.8457	1130	0.2863
0.5471	0.9844	G	A	-0.016	0.006	0.01388	1130	0.4529
0.3268	0.9971	T	A	0.014	0.007	0.04286	1130	0.3268
0.5519	0.97	G	A	-0.015	0.006	0.02337	1130	0.4481
0.4105	0.9897	T	C	-0.004	0.006	0.5204	1130	0.4105
0.6729	0.9902	G	A	-0.013	0.007	0.05273	1130	0.3271
0.3273	0.9945	G	A	0.014	0.007	0.04245	1130	0.3273
0.3266	0.9985	G	A	0.014	0.007	0.04302	1130	0.3266
0.3269	0.9964	G	C	0.014	0.007	0.04274	1130	0.3269

0.0334	0.8181	G	C	-0.02	0.019	0.291	1130	0.0334
0.3266	0.9981	G	A	0.014	0.007	0.04295	1130	0.3266
0.6877	0.7822	G	C	-0.02	0.008	0.007989	1130	0.3123
0.8767	0.9466	T	C	0.022	0.01	0.02675	1130	0.1233
0.6227	0.7968	T	C	0.009	0.007	0.2111	1130	0.3773

KORA3

af_kora3	oevar_kora3	coded_kora3	noncoded_kora3	beta_kora3	se_kora3	p_kora3	n_kora3	maf_kora3
0.4954	0.9951	A	G	0.004	0.005	0.4389	1581	0.4954
0.4951	0.9981	G	C	0.004	0.005	0.4403	1581	0.4951
0.4182	0.8779	G	C	-0.008	0.005	0.1112	1581	0.4182
0.4184	0.8806	G	A	-0.008	0.005	0.1141	1581	0.4184
0.4425	0.9409	T	C	0.008	0.005	0.07597	1581	0.4425
0.444	0.9372	A	G	0.008	0.005	0.07985	1581	0.444
0.4438	0.9376	A	G	0.008	0.005	0.07853	1581	0.4438
0.4451	0.9356	A	T	0.008	0.005	0.08146	1581	0.4451
0.4453	0.9355	T	C	0.008	0.005	0.08426	1581	0.4453
0.4441	0.9244	T	C	0.007	0.005	0.1118	1581	0.4441
0.4442	0.9245	A	C	0.007	0.005	0.1185	1581	0.4442
0.425	0.9981	C	T	0.004	0.005	0.3823	1581	0.425
0.6258	0.9563	G	A	0.005	0.005	0.2782	1581	0.3742
0.4454	0.9354	A	G	0.008	0.005	0.08874	1581	0.4454
0.4455	0.9353	A	G	0.008	0.005	0.09687	1581	0.4455
0.4461	0.9352	G	C	0.008	0.005	0.1025	1581	0.4461
0.4447	0.9588	A	C	0.008	0.005	0.08662	1581	0.4447
0.4445	0.9598	T	A	0.008	0.005	0.0852	1581	0.4445
0.4497	0.941	T	C	0.007	0.005	0.1487	1581	0.4497
0.4489	0.9386	T	C	0.007	0.005	0.1331	1581	0.4489
0.4617	0.9756	A	G	0.006	0.005	0.2142	1581	0.4617
0.49	0.9991	C	T	0.006	0.005	0.1817	1581	0.49
0.4913	0.9867	G	T	0.006	0.005	0.207	1581	0.4913
0.43	1	G	A	0.005	0.005	0.3124	1581	0.43
0.4423	0.1112	A	G	-0.025	0.013	0.06518	1581	0.4423
0.3007	0.9392	C	T	0.001	0.005	0.9182	1581	0.3007
0.6435	0.6868	G	A	0.006	0.006	0.2765	1581	0.3565
0.555	0.9988	A	T	0.004	0.005	0.4335	1581	0.445
0.555	0.9982	C	G	0.004	0.005	0.4331	1581	0.445
0.4654	0.9488	T	A	0.007	0.005	0.1388	1581	0.4654
0.41	1	C	T	0.005	0.005	0.3137	1581	0.41
0.4105	0.995	G	T	0.005	0.005	0.2615	1581	0.4105
0.415	1	G	A	0.005	0.005	0.3164	1581	0.415
0.4441	0.8719	G	C	0.006	0.005	0.2665	1581	0.4441
0.4099	0.9994	T	A	0.005	0.005	0.2679	1581	0.4099
0.4158	0.9944	C	T	0.005	0.005	0.2592	1581	0.4158
0.4098	0.9993	A	G	0.005	0.005	0.2733	1581	0.4098
0.56	1	T	A	0.005	0.005	0.3361	1581	0.44
0.4651	0.9993	C	T	0.007	0.005	0.1321	1581	0.4651
0.4051	0.9522	G	T	0.005	0.005	0.2747	1581	0.4051
0.458	0.9936	A	C	0.008	0.005	0.06635	1581	0.458
0.4655	0.996	A	G	0.006	0.005	0.1711	1581	0.4655
0.5602	0.9779	G	T	0.004	0.005	0.3505	1581	0.4398
0.5601	0.9905	A	G	0.004	0.005	0.3451	1581	0.4399
0.4719	0.9566	C	T	0.007	0.005	0.1544	1581	0.4719
0.5607	0.9529	G	T	0.004	0.005	0.4114	1581	0.4393

0.4758	0.9924	G	C	0.006	0.005	0.1811	1581	0.4758
0.4751	0.9986	C	A	0.006	0.005	0.1797	1581	0.4751
0.4142	0.9773	G	A	0.005	0.005	0.2667	1581	0.4142
0.6176	0.9461	C	G	0.007	0.005	0.1606	1581	0.3824
0.4769	0.9844	C	T	0.006	0.005	0.1927	1581	0.4769
0.4771	0.9837	T	G	0.006	0.005	0.1979	1581	0.4771
0.475	1	A	G	0.006	0.005	0.1745	1581	0.475
0.475	0.9999	A	G	0.006	0.005	0.1746	1581	0.475
0.6269	0.8565	T	A	0.008	0.005	0.1444	1581	0.3731
0.4053	0.9504	C	T	0.005	0.005	0.2847	1581	0.4053
0.4177	0.9838	C	T	0.005	0.005	0.2588	1581	0.4177
0.6345	0.9136	C	T	-0.001	0.005	0.8149	1581	0.3655
0.109	0.9295	G	T	0.002	0.007	0.7614	1581	0.109
0.4199	0.8634	G	A	0.006	0.005	0.2349	1581	0.4199
0.1102	0.9375	C	A	0.002	0.007	0.7309	1581	0.1102
0.4753	0.9868	A	G	0.009	0.005	0.05197	1581	0.4753
0.4751	0.9882	A	T	0.009	0.005	0.04934	1581	0.4751
0.5408	0.9937	A	G	0.005	0.005	0.3143	1581	0.4592
0.4181	0.9832	C	T	0.006	0.005	0.2318	1581	0.4181
0.1459	0.974	A	G	0.003	0.006	0.5932	1581	0.1459
0.6363	0.9301	T	C	-0.001	0.005	0.8777	1581	0.3637
0.54	1	T	A	0.005	0.005	0.3146	1581	0.46
0.111	0.9435	C	A	0.003	0.007	0.6864	1581	0.111
0.4743	0.9944	C	G	0.009	0.005	0.04601	1581	0.4743
0.4744	0.9961	G	C	0.009	0.005	0.05242	1581	0.4744
0.4206	0.9896	C	T	0.006	0.005	0.2316	1581	0.4206
0.6384	0.9128	G	A	-0.001	0.005	0.8136	1581	0.3616
0.12	0.8606	G	A	0.001	0.007	0.8427	1581	0.12
0.4197	0.9886	G	C	0.006	0.005	0.2318	1581	0.4197
0.5391	0.9827	G	A	0.006	0.005	0.1957	1581	0.4609
0.5294	0.9497	G	A	0.005	0.005	0.3241	1581	0.4706
0.5415	0.9818	T	C	0.005	0.005	0.2396	1581	0.4585
0.5399	0.9994	G	C	0.005	0.005	0.3147	1581	0.4601
0.5198	0.9164	G	A	0.006	0.005	0.1728	1581	0.4802
0.2877	0.8573	T	C	-0.011	0.006	0.07039	1573	0.2877
0.5398	0.9992	A	G	0.005	0.005	0.3148	1581	0.4602
0.5984	0.9846	A	T	-0.005	0.005	0.3106	1581	0.4016
0.5396	0.9964	A	G	0.006	0.005	0.1867	1581	0.4604
0.3347	0.9827	T	C	0.001	0.005	0.8059	1581	0.3347
0.3964	0.9689	A	G	0.005	0.005	0.3167	1581	0.3964
0.5988	0.9705	A	G	-0.005	0.005	0.3129	1581	0.4012
0.5995	0.9769	A	G	-0.005	0.005	0.3221	1581	0.4005
0.5986	0.9763	C	G	-0.005	0.005	0.314	1581	0.4014
0.0284	0.7839	G	C	-0.018	0.013	0.1826	1581	0.0284
0.5987	0.9824	A	G	-0.005	0.005	0.3163	1581	0.4013
0.3906	0.9635	C	G	0.002	0.005	0.6632	1581	0.3906
0.0946	0.9163	C	T	0.003	0.007	0.7033	1581	0.0946
0.6294	0.9954	T	C	-0.003	0.005	0.4707	1581	0.3706

KORA4

af_kora4	oevar_kora4	coded_kora4	noncoded_kora4	beta_kora4	se_kora4	p_kora4	n_kora4	maf_kora4
0.4963	0.9842	G	A	-0.01	0.005	0.04198	1407	0.4963
0.5032	0.9893	G	C	0.01	0.005	0.04117	1407	0.4968
0.434	0.9476	G	C	-0.007	0.005	0.1405	1407	0.434
0.4344	0.9468	G	A	-0.007	0.005	0.1397	1407	0.4344
0.4854	0.9606	T	C	0.007	0.005	0.1559	1407	0.4854
0.5143	0.9604	G	A	-0.007	0.005	0.154	1407	0.4857
0.5145	0.9605	G	A	-0.007	0.005	0.155	1407	0.4855
0.5142	0.9604	T	A	-0.007	0.005	0.1535	1407	0.4858
0.486	0.9603	T	C	0.007	0.005	0.1523	1407	0.486
0.4839	0.9507	T	C	0.008	0.005	0.126	1407	0.4839
0.516	0.9509	C	A	-0.008	0.005	0.1256	1407	0.484
0.5746	0.9859	T	C	-0.008	0.005	0.1034	1407	0.4254
0.6358	0.9554	G	A	0.009	0.005	0.08848	1407	0.3642
0.5137	0.9604	G	A	-0.007	0.005	0.1504	1407	0.4863
0.5136	0.9605	G	A	-0.007	0.005	0.1496	1407	0.4864
0.4866	0.9605	G	C	0.007	0.005	0.149	1407	0.4866
0.5091	0.977	C	A	-0.008	0.005	0.1272	1407	0.4909
0.4903	0.9786	T	A	0.008	0.005	0.1305	1407	0.4903
0.4878	0.9631	T	C	0.007	0.005	0.1424	1407	0.4878
0.4875	0.9622	T	C	0.007	0.005	0.144	1407	0.4875
0.5025	0.9865	G	A	-0.008	0.005	0.09371	1407	0.4975
0.4903	0.9983	T	C	-0.007	0.005	0.1599	1407	0.4903
0.4894	0.974	T	G	-0.008	0.005	0.1305	1407	0.4894
0.4326	0.9959	G	A	0.007	0.005	0.1695	1407	0.4326
0.5124	0.9101	G	A	0.013	0.005	0.01157	1407	0.4876
0.6771	0.9883	T	C	-0.013	0.005	0.01582	1407	0.3229
0.6448	0.8068	G	A	0.009	0.006	0.1004	1407	0.3552
0.4587	0.9886	T	A	-0.008	0.005	0.09747	1407	0.4587
0.4589	0.9877	G	C	-0.008	0.005	0.09651	1407	0.4589
0.4964	0.9736	T	A	0.007	0.005	0.1798	1407	0.4964
0.5951	0.9899	T	C	-0.006	0.005	0.2174	1407	0.4049
0.5962	0.9931	T	G	-0.006	0.005	0.2119	1407	0.4038
0.4098	0.9973	G	A	0.005	0.005	0.2981	1407	0.4098
0.4422	0.8914	G	C	0.004	0.005	0.4866	1407	0.4422
0.4023	0.9863	T	A	0.006	0.005	0.2277	1407	0.4023
0.5898	0.994	T	C	-0.005	0.005	0.3145	1407	0.4102
0.5981	0.9864	G	A	-0.006	0.005	0.2226	1407	0.4019
0.5481	0.9949	T	A	0.007	0.005	0.1559	1407	0.4519
0.4932	0.9955	T	C	-0.007	0.005	0.1413	1407	0.4932
0.5977	0.9649	T	G	-0.005	0.005	0.333	1407	0.4023
0.5023	0.9984	C	A	-0.007	0.005	0.1735	1407	0.4977
0.4926	0.9748	G	A	-0.007	0.005	0.1411	1407	0.4926
0.4542	0.9815	T	G	-0.007	0.005	0.1546	1407	0.4542
0.4541	0.9837	G	A	-0.007	0.005	0.1546	1407	0.4541
0.492	0.9645	T	C	-0.007	0.005	0.1439	1407	0.492

0.4516	0.968	T	G	-0.007	0.005	0.1554	1407	0.4516
0.5074	0.9986	G	C	0.008	0.005	0.1113	1407	0.4926
0.5076	0.9996	C	A	0.008	0.005	0.1096	1407	0.4924
0.4059	0.9871	G	A	0.005	0.005	0.2754	1407	0.4059
0.3703	0.9581	G	C	-0.006	0.005	0.2025	1407	0.3703
0.493	0.9976	T	C	-0.008	0.005	0.1147	1407	0.493
0.5068	0.9975	T	G	0.008	0.005	0.116	1407	0.4932
0.4923	1	G	A	-0.008	0.005	0.109	1407	0.4923
0.4923	1	G	A	-0.008	0.005	0.1092	1407	0.4923
0.6339	0.9	T	A	0.005	0.005	0.3857	1407	0.3661
0.598	0.9629	T	C	-0.005	0.005	0.3478	1407	0.402
0.5889	0.9887	T	C	-0.004	0.005	0.3793	1407	0.4111
0.3713	0.9921	T	C	-0.01	0.005	0.06382	1407	0.3713
0.8745	0.9265	T	G	-0.001	0.008	0.8782	1407	0.1255
0.4128	0.8858	G	A	0.004	0.005	0.4376	1407	0.4128
0.1273	0.9355	C	A	0.001	0.008	0.885	1407	0.1273
0.4836	0.9795	G	A	-0.007	0.005	0.1942	1407	0.4836
0.484	0.9817	T	A	-0.007	0.005	0.1974	1407	0.484
0.4854	0.9919	G	A	-0.007	0.005	0.1434	1407	0.4854
0.5893	0.9949	T	C	-0.004	0.005	0.4789	1407	0.4107
0.8078	0.9525	G	A	0.009	0.006	0.1729	1407	0.1922
0.6285	0.9933	T	C	0.01	0.005	0.05774	1407	0.3715
0.5188	0.9911	T	A	0.006	0.005	0.2253	1407	0.4812
0.1286	0.9454	C	A	0.001	0.008	0.8906	1407	0.1286
0.4842	0.9827	G	C	-0.006	0.005	0.199	1407	0.4842
0.5182	0.9968	G	C	0.006	0.005	0.2152	1407	0.4818
0.5881	0.9929	T	C	-0.003	0.005	0.4989	1407	0.4119
0.6287	0.9954	G	A	0.01	0.005	0.05557	1407	0.3713
0.1377	0.8701	G	A	-0.001	0.008	0.9096	1407	0.1377
0.4109	0.9963	G	C	0.003	0.005	0.4895	1407	0.4109
0.4278	0.9961	G	A	0	0.005	0.9269	1407	0.4278
0.429	0.9873	G	A	0	0.005	0.9922	1407	0.429
0.514	0.994	G	A	0.005	0.005	0.2773	1407	0.486
0.5064	0.9642	G	A	0.007	0.005	0.1587	1407	0.4936
0.5144	0.9873	T	C	0.006	0.005	0.1974	1407	0.4856
0.5176	0.9858	G	C	0.006	0.005	0.2085	1407	0.4824
0.5077	0.9896	G	A	0.006	0.005	0.2211	1407	0.4923
0.244	0.8563	T	C	-0.009	0.006	0.1757	1407	0.244
0.4826	0.9857	G	A	-0.006	0.005	0.2048	1407	0.4826
0.3907	0.9894	T	A	0.008	0.005	0.1293	1407	0.3907
0.4864	0.9986	G	A	-0.005	0.005	0.2948	1407	0.4864
0.3784	0.9832	T	C	-0.002	0.005	0.7727	1407	0.3784
0.6123	0.9781	G	A	-0.008	0.005	0.1446	1407	0.3877
0.3909	0.9817	G	A	0.008	0.005	0.1384	1407	0.3909
0.3906	0.9825	G	A	0.008	0.005	0.1246	1407	0.3906
0.3907	0.9871	G	C	0.008	0.005	0.1306	1407	0.3907
0.0299	0.7459	G	C	-0.036	0.017	0.03782	1407	0.0299

0.3906	0.9853	G	A	0.008	0.005	0.1264	1407	0.3906
0.606	0.9575	G	C	-0.008	0.005	0.1173	1407	0.394
0.9004	0.92	T	C	0.006	0.009	0.48	1407	0.0996
0.6364	0.9687	T	C	0.013	0.005	0.01653	1407	0.3636

NESDA

af_nesda	oevar_nesda	coded_nesda	noncoded_nesda	beta_nesda	se_nesda	p_nesda	n_nesda	maf_nesda
0.46540881	0.971693	G	A	-0.00364498	0.00416633	0.381643	1612	0.46540881
0.46541718	0.971819	C	G	-0.00364973	0.00416606	0.380997	1612	0.46541718
0.40694789	1	G	C	-0.00185562	0.0041988	0.658531	1612	0.40694789
0.40888546	0.993784	G	A	-0.00162312	0.00420184	0.699281	1612	0.40888546
0.48632321	0.993663	C	T	-0.00508293	0.00408207	0.213064	1612	0.48632321
0.48636386	0.993944	G	A	-0.00507875	0.00408154	0.213377	1612	0.48636386
0.48634678	0.993825	G	A	-0.00508053	0.00408177	0.213244	1612	0.48634678
0.48636727	0.993971	T	A	-0.00507834	0.00408149	0.213409	1612	0.48636727
0.48637038	0.993992	C	T	-0.00507804	0.00408145	0.213436	1612	0.48637038
0.4918626	0.976989	C	T	-0.00530053	0.0041191	0.198158	1612	0.4918626
0.49186197	0.976979	C	A	-0.00530039	0.00411911	0.198174	1612	0.49186197
0.44843142	0.958411	C	T	0.00386948	0.00424978	0.362541	1612	0.44843142
0.35225456	0.94035	A	G	-0.00144509	0.00440674	0.742962	1612	0.35225456
0.48727498	1	G	A	-0.00497854	0.00407342	0.221632	1612	0.48727498
0.48637193	0.993997	G	A	-0.00507756	0.00408144	0.213476	1612	0.48637193
0.48637193	0.993996	C	G	-0.00507755	0.00408144	0.213479	1612	0.48637193
0.47213801	0.967259	C	A	-0.00601256	0.00413149	0.14559	1612	0.47213801
0.47213832	0.967238	A	T	-0.00601262	0.00413154	0.145587	1612	0.47213832
0.48635859	0.993773	C	T	-0.00507423	0.00408188	0.21383	1612	0.48635859
0.48636542	0.993882	C	T	-0.00507588	0.00408167	0.213657	1612	0.48636542
0.47252111	0.988101	G	A	-0.00543942	0.00409784	0.18438	1612	0.47252111
0.45838509	1	T	C	-0.00404998	0.00408827	0.321864	1612	0.45838509
0.4583978	0.994728	T	G	-0.0041752	0.00409659	0.308109	1612	0.4583978
0.45515289	0.953908	G	A	0.00526904	0.004255	0.215601	1612	0.45515289
0.34088089	1	C	T	-0.00160836	0.00425192	0.705232	1612	0.34088089
0.33293073	0.909614	A	G	-0.00392189	0.00453933	0.387591	1612	0.33293073
0.45169259	0.954399	T	A	-0.00333529	0.0041874	0.425737	1612	0.45169259
0.4517013	0.954508	G	C	-0.00333348	0.0041872	0.425962	1612	0.4517013
0.4841852	0.990073	A	T	-0.00548842	0.00408054	0.178619	1612	0.4841852
0.43854749	1	C	T	0.00388796	0.00418064	0.352375	1612	0.43854749
0.43798209	0.994125	G	T	0.00407638	0.00418374	0.329892	1612	0.43798209
0.45217391	1	G	A	0.00398595	0.00418865	0.341297	1612	0.45217391
0.48005954	0.94368	G	C	0.0050186	0.00429493	0.242609	1612	0.48005954
0.43822112	0.993433	T	A	0.00406778	0.00418317	0.330845	1612	0.43822112
0.451313	0.992712	C	T	0.00379446	0.00419757	0.366013	1612	0.451313
0.43814059	0.995043	A	G	0.00405289	0.00418101	0.332372	1612	0.43814059
0.44482658	0.950497	A	T	-0.0047643	0.00420553	0.257272	1612	0.44482658
0.45934202	1	T	C	-0.00593271	0.00405413	0.143365	1612	0.45934202
0.43935918	0.96884	G	T	0.00341889	0.00427464	0.423821	1612	0.43935918
0.46985699	0.970904	C	A	-0.00640033	0.00411615	0.119965	1612	0.46985699
0.47001533	0.992119	G	A	-0.00546821	0.00408401	0.180588	1612	0.47001533
0.44469851	0.950471	T	G	-0.00478357	0.00420547	0.255338	1612	0.44469851
0.44473287	0.950452	G	A	-0.00477857	0.00420554	0.255846	1612	0.44473287
0.46044933	0.981094	T	C	-0.00460696	0.00411272	0.262634	1612	0.46044933
0.44395875	0.952486	T	G	-0.0047709	0.00420301	0.256324	1612	0.44395875
0.4698534	0.992101	C	G	-0.00550222	0.00408325	0.177821	1612	0.4698534

0.46985311	0.992105	A	C	-0.00550205	0.00408324	0.177833	1612	0.46985311
0.43662996	0.991855	G	A	0.00418714	0.00419209	0.317881	1612	0.43662996
0.35714973	0.946334	G	C	-0.00254274	0.00439359	0.562762	1612	0.35714973
0.46985433	0.992086	T	C	-0.00550278	0.00408328	0.177772	1612	0.46985433
0.46985464	0.992077	G	T	-0.00550311	0.0040833	0.177755	1612	0.46985464
0.4698528	0.992112	G	A	-0.0055018	0.00408323	0.177842	1612	0.4698528
0.4698528	0.992112	G	A	-0.00550179	0.00408323	0.177843	1612	0.4698528
0.34627916	0.940473	A	T	-0.00263962	0.00443128	0.551383	1612	0.34627916
0.43846805	0.969096	C	T	0.00327542	0.00427481	0.443551	1612	0.43846805
0.44940022	0.994577	C	T	0.00370627	0.0041993	0.37746	1612	0.44940022
0.36339533	0.994703	T	C	0.00188323	0.00420029	0.653896	1612	0.36339533
0.12425786	0.980932	G	T	-0.0087608	0.00604846	0.147495	1612	0.12425786
0.46356855	0.985998	G	A	0.00572466	0.00419897	0.172774	1612	0.46356855
0.12280702	1	C	A	-0.00819891	0.0060495	0.175321	1612	0.12280702
0.45866375	0.985546	G	A	-0.00613478	0.00410022	0.134602	1612	0.45866375
0.45866375	0.985545	T	A	-0.00613476	0.00410023	0.134602	1612	0.45866375
0.46227885	0.990477	G	A	-0.00344849	0.00409752	0.40001	1612	0.46227885
0.44619693	0.986835	C	T	0.00344988	0.00421984	0.413624	1612	0.44619693
0.17207379	1	A	G	-0.0107712	0.00528017	0.0413563	1612	0.17207379
0.36379134	0.994318	C	T	0.00190236	0.00420126	0.650693	1612	0.36379134
0.46179805	0.98697	A	T	-0.00352669	0.00410336	0.390089	1612	0.46179805
0.12310367	0.970818	C	A	-0.00820939	0.0060995	0.17833	1612	0.12310367
0.45866375	0.985531	G	C	-0.00613476	0.00410025	0.134605	1612	0.45866375
0.45866437	0.985487	C	G	-0.00613489	0.00410034	0.134607	1612	0.45866437
0.44611104	0.982433	C	T	0.00346548	0.00422597	0.412187	1612	0.44611104
0.36290416	0.983552	A	G	0.00193309	0.00422706	0.647443	1612	0.36290416
0.13412024	0.874686	G	A	-0.0117459	0.00619443	0.0579326	1612	0.13412024
0.4461383	0.983722	G	C	0.00346086	0.00422425	0.412626	1612	0.4461383
0.4185047	0.996382	G	A	-0.00353274	0.00416261	0.39606	1612	0.4185047
0.41820724	0.996293	G	A	-0.00350373	0.00416633	0.400369	1612	0.41820724
0.46702725	0.984582	A	G	-0.00305485	0.00410718	0.457004	1612	0.46702725
0.47423418	0.968306	A	G	-0.0029966	0.00414103	0.469294	1612	0.47423418
0.47041985	1	C	T	-0.00300003	0.00416892	0.47176	1612	0.47041985
0.46196958	0.989026	C	G	-0.00349078	0.00409975	0.394515	1612	0.46196958
0.4772168	0.977349	A	G	-0.00235479	0.00415075	0.570501	1612	0.4772168
0.26330798	1	T	C	-0.00261836	0.00464875	0.573273	1612	0.26330798
0.46206948	0.990479	G	A	-0.00346888	0.00409716	0.397189	1612	0.46206948
0.42324425	1	T	A	0.00365438	0.004175	0.38141	1612	0.42324425
0.46301887	0.969831	G	A	-0.00299547	0.00413494	0.468805	1612	0.46301887
0.4016451	0.918573	T	C	-0.00307281	0.00424024	0.468645	1612	0.4016451
0.4222208	0.984295	A	G	0.00337048	0.00420717	0.423053	1612	0.4222208
0.42333193	0.99483	G	A	0.00346654	0.00418183	0.407131	1612	0.42333193
0.42316344	0.995808	G	A	0.00345157	0.00417921	0.408869	1612	0.42316344
0.42326828	0.995508	G	C	0.00345981	0.00418018	0.407861	1612	0.42326828
0.42314674	0.996062	G	A	0.0034434	0.00417877	0.409931	1612	0.42314674
0.40027978	0.842307	C	G	0.00495077	0.00451507	0.272863	1612	0.40027978
0.11309117	0.94838	C	T	-0.00937341	0.00643155	0.14501	1612	0.11309117
0.35357125	0.811326	C	T	-0.00364625	0.00463303	0.431275	1612	0.35357125

NBS

af_nij	oevar_nij	coded_nij	noncoded_nij	beta_nij	se_nij	p_nij	n_nij	maf_nij
0.45825427	0.984754	G	A	0.00148322	0.00617125	0.810248	527	0.45825427
0.54174573	0.984789	G	C	-0.00147869	0.00617091	0.810431	527	0.458254
0.41340782	0.996125	G	C	0.00483875	0.00624649	0.438324	537	0.41340782
0.41340782	0.996218	G	A	0.00482706	0.00624575	0.439646	537	0.41340782
0.51301115	0.995284	T	C	0.00036301	0.00637473	0.953903	538	0.486989
0.48618785	1	G	A	-0.00074251	0.0063831	0.907396	543	0.48618785
0.48698885	0.995284	G	A	-0.00036301	0.00637473	0.955236	538	0.48698885
0.48698885	0.995284	T	A	-0.00036347	0.0063746	0.954804	538	0.48698885
0.51301115	0.995273	T	C	0.00036069	0.0063745	0.954241	538	0.486989
0.50194553	0.947626	T	C	0.00075124	0.00653716	0.906068	514	0.498054
0.49805447	0.947534	C	A	-0.00074428	0.00653734	0.912869	514	0.49805447
0.53240741	0.997515	T	C	0.00092474	0.00592598	0.876032	540	0.467593
0.66385768	0.996208	G	A	0.00163142	0.00656026	0.805108	534	0.336142
0.48698885	0.995239	G	A	-0.00035673	0.00637449	0.955478	538	0.48698885
0.48698885	0.995239	G	A	-0.00035673	0.00637449	0.955478	538	0.48698885
0.51301115	0.995239	G	C	0.00035673	0.00637449	0.955259	538	0.486989
0.4766537	0.957483	C	A	0.00010434	0.00649609	0.98282	514	0.4766537
0.5233463	0.957472	T	A	-0.00010657	0.00649598	0.991958	514	0.476654
0.51301115	0.994788	T	C	0.00033603	0.00637527	0.958667	538	0.486989
0.51301115	0.994902	T	C	0.00034368	0.00637569	0.956898	538	0.486989
0.46993988	0.957617	G	A	9.25E-05	0.00645856	0.98815	499	0.46993988
0.45664207	0.998999	T	C	0.00186096	0.00620334	0.764183	542	0.45664207
0.45488029	1	T	G	0.00218744	0.00621524	0.724878	543	0.45488029
0.47180451	0.99634	G	A	4.01E-05	0.00592141	0.994163	532	0.47180451
0.49722736	1	G	A	0.00194942	0.00594751	0.743085	541	0.49722736
0.66795367	0.975209	T	C	-0.0021921	0.00672948	0.744932	518	0.332046
0.67625232	1	G	A	0.00314993	0.00662684	0.634552	539	0.323748
0.43761639	0.997274	T	A	0.00059236	0.00598706	0.921134	537	0.43761639
0.43866171	0.997315	G	C	0.00059079	0.00598684	0.921422	538	0.43866171
0.5259481	0.95932	T	A	0.00032921	0.00644233	0.960016	501	0.474052
0.54704797	0.998527	T	C	0.00294295	0.00592023	0.620039	542	0.452952
0.546875	0.999291	T	G	0.00295493	0.00593417	0.619443	544	0.453125
0.46494465	0.999212	G	A	-0.00290239	0.00592788	0.624782	542	0.46494465
0.49238095	0.988323	G	C	-2.49E-05	0.00595735	0.99868	525	0.49238095
0.45378928	0.99801	T	A	-0.00297545	0.00593843	0.615944	541	0.45378928
0.53406998	1	T	C	0.00254301	0.00594724	0.668946	543	0.46593
0.546875	0.998579	G	A	0.00296675	0.00593653	0.618494	544	0.453125
0.56685499	0.996427	T	A	0.00031659	0.00597149	0.957718	531	0.433145
0.45664207	0.998979	T	C	0.00185821	0.0062034	0.764807	542	0.45664207
0.53531599	0.999323	T	G	0.00287307	0.0059332	0.628132	538	0.464684
0.47	0.947314	C	A	9.83E-05	0.00648741	0.984125	500	0.47
0.45664207	0.998999	G	A	0.00186088	0.00620321	0.764187	542	0.45664207
0.43301887	0.996282	T	G	-0.00028677	0.00597232	0.961174	530	0.43301887
0.43301887	0.99634	G	A	-0.00029182	0.00597201	0.960806	530	0.43301887
0.45664207	0.998999	T	C	0.00186088	0.00620321	0.764187	542	0.45664207

0.43045113	0.996451	T	G	-0.0002283	0.00596214	0.969354	532	0.43045113
0.54335793	0.998578	G	C	-0.00184225	0.00620522	0.766552	542	0.456642
0.54335793	0.998578	C	A	-0.00184225	0.00620522	0.766552	542	0.456642
0.44944853	1	G	A	-0.00263611	0.00587942	0.653892	544	0.44944853
0.34116541	0.995762	G	C	-0.00377765	0.00642383	0.55642	532	0.34116541
0.45664207	0.99849	T	C	0.00183331	0.00620566	0.767669	542	0.45664207
0.54335793	0.998359	T	G	-0.00182445	0.00620646	0.769078	542	0.456642
0.45664207	0.998679	G	A	0.00184825	0.00620477	0.765798	542	0.45664207
0.45664207	0.998679	G	A	0.00184825	0.00620477	0.765798	542	0.45664207
0.66698473	0.985361	T	A	0.00368344	0.00655837	0.571905	524	0.333015
0.53531599	0.998973	T	C	0.00285741	0.00593197	0.629788	538	0.464684
0.53531599	0.997612	T	C	0.00273246	0.00590769	0.643726	538	0.464684
0.36832061	0.976354	T	C	-0.00087591	0.00652769	0.894169	524	0.36832061
0.87775735	1	T	G	0.0118556	0.009584	0.216081	544	0.122243
0.48154982	1	G	A	-0.00187878	0.00590897	0.75052	542	0.48154982
0.12224265	0.997486	C	A	-0.0118691	0.00958849	0.215774	544	0.12224265
0.46319018	0.941521	G	A	-0.00059776	0.00647569	0.929568	489	0.46319018
0.46319018	0.941521	T	A	-0.00059776	0.00647569	0.929568	489	0.46319018
0.45404412	0.999295	G	A	0.0026037	0.0060045	0.664585	544	0.45404412
0.5380334	0.997256	T	C	0.0028188	0.00589522	0.631544	539	0.461967
0.84579439	0.987608	G	A	0.0188647	0.00861777	0.0285009	535	0.154206
0.63167939	0.97581	T	C	0.00084733	0.00652806	0.900391	524	0.368321
0.5494403	0.997345	T	A	-0.00270701	0.00601306	0.651753	536	0.45056
0.12224265	0.994657	C	A	-0.0119302	0.00959704	0.213607	544	0.12224265
0.46319018	0.941521	G	C	-0.00059776	0.00647569	0.929568	489	0.46319018
0.53680982	0.941521	G	C	0.00059776	0.00647569	0.922797	489	0.46319
0.53810409	0.996338	T	C	0.00281686	0.00589785	0.633258	538	0.461896
0.62065637	0.970023	G	A	0.00043608	0.00653312	0.949437	518	0.379344
0.12524272	0.919777	G	A	-0.0124523	0.00969249	0.198646	515	0.12524272
0.46189591	0.996828	G	C	-0.00280012	0.00589642	0.634815	538	0.46189591
0.39430147	1	G	A	-0.00785967	0.00627194	0.210152	544	0.39430147
0.39430147	1	G	A	-0.00785967	0.00627194	0.210152	544	0.39430147
0.54267161	0.996842	G	A	-0.00242388	0.00593935	0.681624	539	0.457328
0.55606618	1	G	A	-0.00282291	0.00601544	0.638871	544	0.443934
0.5464684	0.997615	T	C	-0.00237992	0.00596322	0.689756	538	0.453532
0.54805915	0.997992	G	C	-0.00267345	0.00601031	0.656114	541	0.451941
0.53584559	1	G	A	-0.00470841	0.00591635	0.42613	544	0.464154
0.24372385	0.912671	T	C	-0.00026907	0.0074579	0.970053	478	0.24372385
0.45395948	0.998344	G	A	0.00264858	0.00600832	0.659307	543	0.45395948
0.43081181	0.999772	T	A	-0.00424698	0.00598404	0.47788	542	0.43081181
0.4447619	0.986338	G	A	0.00289283	0.00601579	0.629495	525	0.4447619
0.37947269	0.994339	T	C	0.00202062	0.00618875	0.744048	531	0.37947269
0.56754221	0.996371	G	A	0.00445353	0.00598412	0.456832	533	0.432458
0.43081181	0.999542	G	A	-0.00426635	0.0059841	0.475999	542	0.43081181
0.43186004	1	G	A	-0.00455285	0.00600495	0.448341	543	0.43186004
0.43081181	0.99972	G	C	-0.00425152	0.00598418	0.477436	542	0.43081181
0.01886792	0.88365	G	C	-0.00787422	0.0207814	0.704733	530	0.01886792

0.43081181	0.999772	G	A	-0.00424698	0.00598404	0.47788	542	0.43081181
0.64340102	0.84269	G	C	0.00213481	0.00668969	0.751868	394	0.356599
0.89408397	0.985945	T	C	0.0119016	0.00983129	0.227288	524	0.105916
0.65508021	0.858921	T	C	-0.00393181	0.00675505	0.551012	374	0.34492

Rotterdam Study

af_rott	oevar_rott	coded_rott	noncoded_rott	beta_rott	se_rott	p_rott	n_rott	maf_rott
0.464583	0.975397	G	A	-0.0062218	0.00384133	0.10527406	5169	0.464583
0.464599	0.975617	C	G	-0.00624191	0.00384075	0.10410128	5169	0.464599
0.584212	0.995835	C	G	0.00356825	0.00387065	0.35652094	5169	0.415788
0.584053	0.996161	A	G	0.00355522	0.0038699	0.35818694	5169	0.415947
0.483684	0.989937	C	T	-0.00779184	0.00381492	0.04110721	5169	0.483684
0.488885	0.999582	G	A	-0.0072542	0.0037928	0.05579221	5169	0.488885
0.488682	0.998733	G	A	-0.00727865	0.00379449	0.05508008	5169	0.488682
0.488534	0.998581	T	A	-0.00728178	0.00379407	0.05494844	5169	0.488534
0.487831	0.998822	C	T	-0.00731994	0.00379215	0.05356769	5169	0.487831
0.49373	0.976933	C	T	-0.00739168	0.00383192	0.05373057	5169	0.49373
0.493713	0.976805	C	A	-0.00738675	0.00383217	0.05390625	5169	0.493713
0.55782	0.973453	T	C	-0.00070363	0.00386787	0.85562132	5169	0.44218
0.351445	0.957304	A	G	-0.00296968	0.00407763	0.46636396	5169	0.351445
0.487799	0.998606	G	A	-0.00731434	0.00379253	0.053775	5169	0.487799
0.487769	0.998387	G	A	-0.00730683	0.00379289	0.05404431	5169	0.487769
0.487736	0.998172	C	G	-0.00730062	0.00379328	0.05427373	5169	0.487736
0.477094	0.921396	C	A	-0.00766627	0.00395385	0.05250608	5169	0.477094
0.477081	0.920495	A	T	-0.00767008	0.00395568	0.05249792	5169	0.477081
0.487483	0.996492	C	T	-0.00724668	0.00379638	0.05627868	5169	0.487483
0.487563	0.997015	C	T	-0.00726394	0.00379538	0.05563004	5169	0.487563
0.473009	0.979077	G	A	-0.00733977	0.00384077	0.05599856	5169	0.473009
0.458239	0.998612	T	C	-0.00681728	0.0038045	0.07313851	5169	0.458239
0.458317	0.998812	T	G	-0.00684281	0.003804	0.07203236	5169	0.458317
0.553979	0.971172	A	G	-0.00118146	0.00386852	0.76001472	5169	0.446021
0.517925	0.993687	G	A	0.00570184	0.0038569	0.13928053	5169	0.482075
0.665569	0.999744	T	C	-0.00427619	0.00399781	0.28471819	5169	0.334431
0.333348	0.929186	A	G	-0.0042799	0.00419945	0.30805882	5169	0.333348
0.452536	0.972188	T	A	0.00180156	0.00388385	0.64268789	5169	0.452536
0.452486	0.972506	G	C	0.00178755	0.00388335	0.64523238	5169	0.452486
0.482296	0.983227	A	T	-0.00731517	0.00382446	0.05577781	5169	0.482296
0.568183	0.996129	T	C	-0.00111709	0.00383483	0.77077764	5169	0.431817
0.56793	0.998727	T	G	-0.00121311	0.00382897	0.75133099	5169	0.43207
0.55663	0.996553	A	G	-0.00158159	0.00383516	0.67999541	5169	0.44337
0.532004	0.955	C	G	-0.00183377	0.00390904	0.6389287	5169	0.467996
0.569979	0.997478	A	T	-0.00068313	0.00383183	0.85847811	5169	0.430021
0.556507	0.999471	T	C	-0.00165538	0.00382901	0.66544716	5169	0.443493
0.56989	0.997682	G	A	-0.00071414	0.00383143	0.8521104	5169	0.43011
0.449666	0.966873	A	T	0.00146798	0.00389673	0.70632839	5169	0.449666
0.465654	0.979408	T	C	-0.00707687	0.0038465	0.06578659	5169	0.465654
0.556593	0.998924	T	G	-0.00170254	0.00383058	0.65665044	5169	0.443407
0.472174	0.893183	C	A	-0.00770004	0.00401641	0.05521537	5169	0.472174
0.465393	0.977333	G	A	-0.00709026	0.00385053	0.06555914	5169	0.465393
0.449776	0.966551	T	G	0.00149872	0.00389725	0.70051179	5169	0.449776
0.449743	0.966644	G	A	0.00149014	0.00389709	0.70213246	5169	0.449743
0.464194	0.970138	T	C	-0.00713989	0.00386462	0.06466712	5169	0.464194

0.449302	0.966437	T	G	0.00141958	0.00389858	0.71571342	5169	0.449302
0.466893	0.990165	C	G	-0.0070277	0.0038254	0.06618455	5169	0.466893
0.466847	0.990406	A	C	-0.00701556	0.00382507	0.06663077	5169	0.466847
0.568967	0.998792	A	G	-0.00143739	0.00383078	0.7074434	5169	0.431033
0.360914	0.961208	G	C	-0.00234904	0.00404595	0.56144759	5169	0.360914
0.467049	0.989381	T	C	-0.00706694	0.00382651	0.0647628	5169	0.467049
0.467091	0.989165	G	T	-0.00707808	0.00382681	0.06436343	5169	0.467091
0.466824	0.990518	G	A	-0.00700985	0.0038249	0.06684108	5169	0.466824
0.466791	0.990697	G	A	-0.00700191	0.00382467	0.0671332	5169	0.466791
0.350803	0.945963	A	T	-0.00285839	0.00412229	0.48798554	5169	0.350803
0.557488	0.996489	T	C	-0.00157952	0.00383398	0.68029851	5169	0.442512
0.557864	0.996724	T	C	-0.00164435	0.00383449	0.66798854	5169	0.442136
0.378201	0.998692	T	C	-0.00536299	0.00391734	0.17094346	5169	0.378201
0.871735	0.999992	T	G	0.00775369	0.00570341	0.17394864	5169	0.128265
0.545898	0.998405	A	G	-0.00136893	0.00383203	0.72086753	5169	0.454102
0.871734	0.999992	A	C	0.0077538	0.00570342	0.17394282	5169	0.128266
0.462009	0.888256	G	A	-0.00803149	0.00402046	0.04575423	5169	0.462009
0.461965	0.883765	T	A	-0.00805124	0.00403037	0.04575558	5169	0.461965
0.462063	0.998608	G	A	0.00115867	0.003836	0.76256891	5169	0.462063
0.557194	0.987588	T	C	-0.00241727	0.00383927	0.52887326	5169	0.442806
0.823665	0.999836	G	A	0.00854046	0.00498245	0.08649262	5169	0.176335
0.377958	0.999155	C	T	-0.00509551	0.00391777	0.19334017	5169	0.377958
0.464761	0.995693	A	T	0.00177652	0.00383517	0.64314771	5169	0.464761
0.871734	0.99999	A	C	0.00775358	0.00570341	0.17395446	5169	0.128266
0.461952	0.882334	G	C	-0.00805321	0.00403347	0.04586897	5169	0.461952
0.461895	0.877154	C	G	-0.00807279	0.00404502	0.04596218	5169	0.461895
0.557249	0.987158	T	C	-0.002381	0.00383979	0.53512933	5169	0.442751
0.377229	0.994349	A	G	-0.00528464	0.00392814	0.17847225	5169	0.377229
0.860267	0.9176	A	G	0.00770625	0.00574192	0.17951556	5169	0.139733
0.557215	0.987397	C	G	-0.00240511	0.0038395	0.53097356	5169	0.442785
0.418045	0.999697	G	A	-0.00250572	0.00385213	0.51531235	5169	0.418045
0.418678	0.999895	G	A	-0.00251322	0.00385025	0.51385107	5169	0.418678
0.467949	0.988571	A	G	0.00081232	0.00383591	0.83225743	5169	0.467949
0.46185	0.999259	A	G	0.00109171	0.00383537	0.775876	5169	0.46185
0.463434	0.995756	C	T	0.00117277	0.00383622	0.7597825	5169	0.463434
0.464445	0.997807	C	G	0.00170291	0.00383184	0.65668702	5169	0.464445
0.477854	0.996228	A	G	-0.00051896	0.00381664	0.89182215	5169	0.477854
0.743318	0.999471	C	T	-0.00021943	0.00431823	0.95946623	5169	0.256682
0.464363	0.998439	G	A	0.00168212	0.00383087	0.66053365	5169	0.464363
0.415818	0.997987	T	A	0.0009358	0.00385118	0.80797623	5169	0.415818
0.463985	0.977887	G	A	0.00085479	0.003862	0.82480009	5169	0.463985
0.603234	0.994256	C	T	0.0114317	0.00389272	0.00332455	5169	0.396766
0.583851	0.993373	G	A	-0.00105243	0.00385864	0.78500758	5169	0.416149
0.416001	0.99687	G	A	0.00095535	0.00385258	0.80411709	5169	0.416001
0.415652	0.999052	G	A	0.00091802	0.00384982	0.81148982	5169	0.415652
0.415868	0.997675	G	C	0.0009411	0.00385158	0.80693112	5169	0.415868
0.970674	0.88263	C	G	0.00456312	0.0118828	0.70091669	5169	0.029326

0.415685	0.998825	G	A	0.00092143	0.00385008	0.81081577	5169	0.415685
0.614657	0.829946	G	C	-0.00166939	0.00427054	0.69581141	5169	0.385343
0.891477	0.98329	T	C	0.0101947	0.00615317	0.09753467	5169	0.108523
0.380196	0.817464	C	T	-0.0085947	0.00432185	0.04673787	5169	0.380196

SHIP

af_ship	oevar_ship	coded_ship	noncoded_ship	beta_ship	se_ship	p_ship	n_ship	maf_ship
0.459686	0.986475	G	A	-0.0109506	0.00759535	0.149372	543	0.459686
0.53622652	0.944519	G	C	0.0113383	0.00775203	0.14357	543	0.463773
0.42077293	0.978388	G	C	-0.0103203	0.00770528	0.180447	543	0.42077293
0.42080884	0.978243	G	A	-0.0103426	0.00770588	0.17954	543	0.42080884
0.51883425	0.973215	T	C	0.0107815	0.00774395	0.163846	543	0.481166
0.48113168	0.973312	G	A	-0.0107783	0.00774331	0.163936	543	0.48113168
0.48114457	0.973268	G	A	-0.0107797	0.00774363	0.163898	543	0.48114457
0.48112615	0.973318	T	A	-0.0107776	0.00774327	0.163964	543	0.48112615
0.51887753	0.973328	T	C	0.0107779	0.00774319	0.163948	543	0.481122
0.51775691	0.969361	T	C	0.0107533	0.00775465	0.165536	543	0.482243
0.48224217	0.969375	C	A	-0.0107543	0.00775452	0.165488	543	0.48224217
0.55870074	0.995808	T	C	-0.0111998	0.00786546	0.154469	543	0.441299
0.63335912	0.986561	G	A	0.00510464	0.00797095	0.521909	543	0.366641
0.48112339	0.973331	G	A	-0.0107772	0.00774318	0.163974	543	0.48112339
0.48111971	0.973348	G	A	-0.010777	0.00774307	0.163976	543	0.48111971
0.51887937	0.973348	G	C	0.0107767	0.00774307	0.163986	543	0.481121
0.47749816	0.991734	C	A	-0.0120101	0.00776407	0.121892	543	0.47749816
0.5224954	0.991775	T	A	0.0120109	0.00776403	0.121865	543	0.477505
0.51894843	0.97358	T	C	0.0107706	0.00774154	0.164142	543	0.481052
0.51891344	0.973463	T	C	0.0107738	0.0077423	0.16406	543	0.481087
0.47296869	0.939647	G	A	-0.0120372	0.00781749	0.123614	543	0.47296869
0.45464088	0.993633	T	C	-0.0120655	0.00765034	0.114768	543	0.45464088
0.44805157	0.983274	T	G	-0.0116322	0.00768985	0.130363	543	0.44805157
0.448907	0.994173	G	A	0.0104555	0.00781975	0.1812	543	0.448907
0.52776519	0.167303	G	A	-0.00196393	0.0185514	0.91569	543	0.472235
0.65126796	0.993325	T	C	-0.0103311	0.00792857	0.192567	543	0.348732
0.64657827	0.947161	G	A	0.00041351	0.00827237	0.960133	543	0.353422
0.46497053	0.995873	T	A	-0.00923825	0.00797478	0.246687	543	0.46497053
0.46497238	0.995857	G	C	-0.00923757	0.0079748	0.246723	543	0.46497238
0.5302081	0.977881	T	A	0.0111221	0.00768731	0.147951	543	0.469792
0.56971363	0.992749	T	C	-0.0132359	0.00785264	0.0918858	543	0.430286
0.57102486	0.993873	T	G	-0.012138	0.00784679	0.121893	543	0.428975
0.44087201	0.996213	G	A	0.0118147	0.00795893	0.137689	543	0.44087201
0.46384807	0.963607	G	C	0.00896057	0.00820423	0.27475	543	0.46384807
0.42659042	0.997072	T	A	0.0134643	0.00786656	0.0869739	543	0.42659042
0.56332689	0.992758	T	C	-0.0106475	0.00793568	0.179684	543	0.436673
0.57294199	0.996012	G	A	-0.0134089	0.00786356	0.0881584	543	0.427058
0.54107827	0.996873	T	A	0.00868325	0.00795221	0.274864	543	0.458922
0.45322007	0.970119	T	C	-0.0117252	0.00772981	0.129298	543	0.45322007
0.57121823	0.960638	T	G	-0.0104068	0.00821462	0.205203	543	0.428782
0.46614641	0.985924	C	A	-0.012706	0.00774329	0.100816	543	0.46614641
0.45753775	0.996347	G	A	-0.0113569	0.0076602	0.138183	543	0.45753775
0.45948435	0.993668	T	G	-0.00858144	0.00800685	0.283827	543	0.45948435
0.45948987	0.995636	G	A	-0.00869235	0.00798029	0.276053	543	0.45948987
0.45456262	0.975495	T	C	-0.0115877	0.00771876	0.133294	543	0.45456262

0.45908195	0.983047	T	G	-0.00764492	0.00814818	0.348123	543	0.45908195
0.54307919	0.997554	G	C	0.0113451	0.00764581	0.137852	543	0.456921
0.54307827	0.997558	C	A	0.0113445	0.00764579	0.137874	543	0.456922
0.4316779	0.988736	G	A	0.012126	0.00785497	0.122654	543	0.4316779
0.36804843	0.982422	G	C	-0.00272745	0.00812059	0.73697	543	0.36804843
0.45691897	0.997553	T	C	-0.0113456	0.00764581	0.137836	543	0.45691897
0.54308195	0.99755	T	G	0.0113462	0.00764582	0.137815	543	0.456918
0.45736924	0.99672	G	A	-0.0113819	0.00765368	0.136984	543	0.45736924
0.45737109	0.996717	G	A	-0.011382	0.00765367	0.136982	543	0.45737109
0.64137293	0.969265	T	A	0.00336041	0.00829215	0.685293	543	0.358627
0.57200092	0.959406	T	C	-0.0106185	0.00822669	0.196796	543	0.427999
0.56235359	0.988854	T	C	-0.0108465	0.00794866	0.172387	543	0.437646
0.37309853	0.997157	T	C	-0.00481856	0.00791766	0.5428	543	0.37309853
0.88165838	0.970473	T	G	-0.00763968	0.0119609	0.523003	543	0.118342
0.44195101	0.975386	G	A	0.00842891	0.00805541	0.295392	543	0.44195101
0.11871315	0.97708	C	A	0.00795742	0.0119383	0.505061	543	0.11871315
0.45435912	0.996356	G	A	-0.0113565	0.00772852	0.141716	543	0.45435912
0.45435912	0.996356	T	A	-0.0113565	0.00772852	0.141716	543	0.45435912
0.48123757	0.991337	G	A	-0.00996079	0.00784058	0.203937	543	0.48123757
0.56348987	0.9926	T	C	-0.0108145	0.00794266	0.173334	543	0.43651
0.83353591	0.973577	G	A	-0.0090938	0.0099117	0.35889	543	0.166464
0.62696961	0.997072	T	C	0.00483154	0.0079174	0.5417	543	0.37303
0.52126703	0.996285	T	A	0.0114748	0.00787762	0.145219	543	0.478733
0.11921	0.988668	C	A	0.00837811	0.0119052	0.481597	543	0.11921
0.45436096	0.996379	G	C	-0.0113562	0.00772845	0.141723	543	0.45436096
0.55253775	0.929136	G	C	0.0137417	0.00802022	0.0866426	543	0.447462
0.56200921	0.993575	T	C	-0.010712	0.00795096	0.177897	543	0.437991
0.62678637	0.979427	G	A	0.004209	0.00797248	0.59754	543	0.373214
0.1309618	0.923353	G	A	0.0137601	0.0118015	0.243628	543	0.1309618
0.43715396	0.995151	G	C	0.0107899	0.00794355	0.174362	543	0.43715396
0.38537569	0.996956	G	A	-0.00981888	0.00795422	0.217045	543	0.38537569
0.38558287	0.993777	G	A	-0.0101803	0.00799874	0.203112	543	0.38558287
0.51660681	0.989133	G	A	0.00908687	0.00786849	0.248155	543	0.483393
0.51045764	0.961637	G	A	0.00965181	0.00809391	0.233074	543	0.489542
0.51874862	0.989337	T	C	0.0101464	0.00785672	0.196554	543	0.481251
0.52031676	0.997197	G	C	0.0112935	0.00788147	0.151882	543	0.479683
0.50969429	0.955598	G	A	0.00978531	0.00790963	0.216036	543	0.490306
0.2269663	0.90447	T	C	-0.00278971	0.00962386	0.771913	543	0.2269663
0.47800092	0.997753	G	A	-0.0110825	0.00785415	0.158234	543	0.47800092
0.42189669	0.978759	T	A	0.0152775	0.00798404	0.0556826	543	0.42189669
0.47768692	0.975452	G	A	-0.010603	0.00789852	0.179467	543	0.47768692
0.34436556	0.986964	T	C	0.00990849	0.00781596	0.204896	543	0.34436556
0.58027256	0.988324	G	A	-0.0130404	0.00791982	0.0996502	543	0.419727
0.42163112	0.995303	G	A	0.0130349	0.0078584	0.0971719	543	0.42163112
0.42154807	0.995632	G	A	0.0130904	0.00785495	0.0956099	543	0.42154807
0.42157735	0.99582	G	C	0.013064	0.00785565	0.09631	543	0.42157735
0.03104144	0.719309	G	C	-0.0221084	0.0273471	0.41884	543	0.03104144

0.42153039	0.995917	G	A	0.0130873	0.00785433	0.0956628	543	0.42153039
0.62990055	0.829774	G	C	-0.00885344	0.00871733	0.309813	543	0.370099
0.89743554	0.95325	T	C	-0.0012671	0.0128217	0.921277	543	0.102564
0.62997882	0.854452	T	C	0.002174	0.00848338	0.797746	543	0.370021

Journal of Gerontology: MEDICAL SCIENCES
Cite journal as: J Gerontol A Biol Sci Med Sci. 2012 May;67A(5):470–479
doi:10.1093/gerona/gls089

Translational Article

Special Issue on Genetics and Aging

Translational

The Search for Longevity and Healthy Aging Genes: Insights From Epidemiological Studies and Samples of Long-Lived Individuals

Joanne M. Murabito,[1,2] Rong Yuan,[3] and Kathryn L. Lunetta[4]

[1]Framingham Heart Study, National Heart, Lung and Blood Institute, Framingham, Massachusetts.
[2]Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Massachusetts.
[3]Aging Center, Jackson Laboratory, Bar Harbor, Maine.
[4]Department of Biostatistics, Boston University School of Public Health, Massachusetts.

Address correspondence to Joanne M. Murabito, MD, ScM, Framingham Heart Study, National Heart, Lung, and Blood Institute, 73 Mount Wayte Avenue, Suite 2, Framingham, MA 01702. Email: murabito@bu.edu

Genetic factors clearly contribute to exceptional longevity and healthy aging in humans, yet the identification of the underlying genes remains a challenge. Longevity is a complex phenotype with modest heritability. Age-related phenotypes with higher heritability may have greater success in gene discovery. Candidate gene and genome-wide association studies (GWAS) for longevity have had only limited success to date. The Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium conducted a meta-analysis of GWAS data for longevity, defined as survival to age 90 years or older, that identified several interesting associations but none achieved genome-wide significance. A recent GWAS of longevity conducted in the Leiden Longevity Study identified the *ApoE* E4 isoform as deleterious to longevity that was confirmed in an independent GWAS of long-lived individuals of German descent. Notably, no other genetic loci for longevity have been identified in these GWAS. To examine the conserved genetic mechanisms between the mouse and humans for life span, we mapped the top Cohorts for Heart and Aging Research in Genomic Epidemiology GWAS associations for longevity to the mouse chromosomal map and noted that eight of the ten top human associations were located within a previously reported mouse life-span quantitative trait loci. This work suggests that the mouse and human may share mechanisms leading to aging and that the mouse model may help speed the understanding of how genes identified in humans affect the biology of aging. We expect these ongoing collaborations and the translational work with basic scientists to accelerate the identification of genes that delay aging and promote a healthy life span.

Key Words: Longevity—Genetics—Epidemiological studies.

Received August 29, 2011; Accepted February 23, 2012

Decision Editor: Luigi Ferrucci, MD, PhD

IT is well established that both genetic factors and health-related behaviors influence survival to old age and survival to old age in good health. Longitudinal cohort studies demonstrate that lower levels of cardiovascular risk factors measured in midlife or early older years predict survival and healthy survival to 85 years of age (1,2) and beyond (3,4). Longevity has been observed to cluster within families so that parents and siblings of centenarians have a greater likelihood of attaining advanced age (5–7), and offspring of centenarians appear to have a delay in age-related disease (8,9). Studies of families clustered for longevity both in the United States and Europe (Long Life Family Study and Leiden Longevity Study) have demonstrated that offspring of long-lived participants has more favorable midlife risk factor profiles and less age-related disease (10,11). Similarly in the community-based Framingham Heart Study, adults with at least one parent surviving to old age have lower risk factor levels compared with individuals whose parents died younger and the risk factor advantage persists over time (12). The genetic contribution to longevity and human aging is likely to result from many genes each with modest effects. Some genes will likely affect longevity by increasing susceptibility to age-related disease and early death, whereas other genes are likely to slow the aging process itself leading to a long life. How genetic factors and their interaction with modifiable behavioral and environmental factors contribute to longevity remains unknown.

Longevity and Healthy Aging Phenotypes: Definitions and Heritability

Longevity is often defined as age at death or survival to an exceptional age such as 90 years or older or 100 years or older. Because life expectancy has improved dramatically across birth cohorts since 1900, care must be taken when study designs compare long-lived with younger cohorts. Women live longer than men and make up a larger proportion of the older population especially at exceptional old ages. For example, among the original 5,209 Framingham Heart Study participants with follow-up through 2011, there are 43 centenarian women and only six centenarian men, whereas at study entry (1948–1953), 55% of participants were women. Men are more likely to attain extreme old age escaping common age-related disease, whereas women are more likely to attain 100 after surviving common morbidities (13). While these observations raise the hypothesis that genetic and environmental factors influence the path to longevity differently in men and women, whether genetic factors play a greater or lesser role in men than in women is an area of debate (3,14). In a study of centenarians (100–104 years), semisupercentenarians (105–109 years), and supercentenaians (110–119 years), there was a progressive delay in the onset of age-related disease and onset of physical and cognitive function impairment with increasing age (15). Whether genes that influence survival to these extreme ages also play a role in survival to age older than 90 years is unknown.

The genetic contribution to longevity (age at death) has been estimated using both large twin registries and population-based samples (Table 1). Most heritability estimates from twin registries range between 20% and 30% (16,17), whereas estimates from population-based samples are slightly lower, ranging from 15% to 25% (18,19), suggesting a significant but modest genetic contribution to the human life span. One study conducted among an ethnically diverse group suggests that genetic influences on life span may vary by ethnicity with heritabilities ranging from a low of 4% for African Americans to 29% and 26% for Caribbean Hispanics and Caucasians, respectively (21). Using data from the GenomeEUtwin project that included more than 20,000 Nordic twins, Hjelmborg and coworkers (42) noted that genetic effects on life span were minimal prior to 60 years of age, but genetic effects on life spans greater than 60 years of age were significant and constant to increasing with advancing age. Starting at about age 60 years, the relative recurrence risk of an individual living past a specified age given that his/her cotwin also lived past that age increased with increasing age cut point in both men and women so that at age 92, the recurrence risk was 4.8 and 1.8 in monozygotic and dizygotic male twins and 2.5 and 1.6 in monozygotic and dizygotic female twins. Notably, recurrence risks similar to men occurred in women at a 5- to 10-year older age perhaps reflecting the longer average longevity in women. Using the Framingham Heart Study cohorts, we explored whether genetic influences on life span increase with achievement of older ages by examining age at death as a dichotomous trait using a liability threshold model adjusting for sex and birth year (20). In contrast to the modest heritability estimate for continuous age at death (16%), heritability of surviving past 65 years and surviving past 85 years was substantial at 36% ($p = 4.2 \times 10^{-10}$) and 40% ($p = 9.0 \times 10^{-10}$), respectively. Thus, genetic effects appear to be greater for survival to more advanced ages. In the Framingham Heart Study cohorts, heritability appears to increase with each 10-year increment in survived age (65 years, 75 years, and 85 years) for men, but not women, again suggesting that genetic effects on aging may be more substantial for men than women (20).

Table 1. Familiality of Aging Phenotypes

Exceptional survival: centenarians	Sibling survival probability
New England Centenarian Study, likelihood of achieving age 100 (6)	Women eightfold; men 17-fold
Okinawa Centenarian Study (7), likelihood of achieving age 90	Women 2.6-fold; men 5.4-fold
Age at death	**Heritability**
Twin registries (16,17)	20%–30%
Old Order Amish (18)	25%
Utah Population Database (19)	15%
Framingham Heart Study (20)	16%
Medicare recipients, New York City (21)	
European ancestry	26%
African American	4%
Caribbean Hispanic	29%
Age at death ≥65 y	
Framingham Heart Study ≥65, ≥75, and ≥85 (20)	36%–40%
Healthy aging and morbidity-free survival	
Male twins at age 70 y (22)	50%
Framingham Heart Study ≥65, ≥75, and ≥85 (20)	20%–25%
Physical function, disability, and self-report	
Danish twins, aged 80 and older, women (23)	34%–47%
Danish twins, aged 75–79, women (23)	15%–34%
Rate of change in functional ability*, (24)	ns
Framingham Heart Study, disability free, aged 75‡	44%
Frailty	
Framingham Heart Study, frailty/prefrailty vs no frailty, aged >60 y‡	19%
Danish twins, cluster analysis approach†, (25)	43%
Handgrip	
Twin studies (26–28)	40%–65%
Long Life Family Study (29)	~40%
Framingham Heart Study‡	38%
Walking speed	
Female twins (30)	16%
Male twins (31)	42%
Framingham Heart Study, usual pace walk‡	38%
Framingham Heart Study, quick walk‡	36%
Long Life Family Study (29)	10%
Bone mineral density (32,33)	50%–70%
Alzheimer's disease (34,35)	58%–79%
Reproductive aging	
Age at menarche (36,37)	50%–80%
Age at natural menopause (38–41)	44%–87%

Note: ns = nonsignificant.

*60% of sample did not participate in 4 year follow-up, half lost to mortality.

†Cluster analysis approach based on mini-mental state examination, activities of daily living, self-reported health status, and handgrip strength.

‡Not previously published.

The longevity phenotype measures overall life span without consideration of health and physical or cognitive function and hence is a very heterogeneous phenotype that may be affected by many environmental and other nongenetic factors. The relative contribution of additive genetic effects may be greater for more homogeneous phenotypes that describe specific aspects of aging and in turn may result in greater success in gene discovery. The heritability of reproductive aging phenotypes is at least 50%, and heritability is even higher for age-related diseases such as osteoporosis (low bone mineral density) and Alzheimer's disease (Table 1). Genetic association studies have been successful in identifying genetic variants for these aging phenotypes and have the potential to uncover new biologic insights into the associated underlying aging processes (43–49).

Epidemiological studies that have followed participants over adulthood and collected a wealth of information in a standardized fashion have been important sources for development of aging-related phenotypes. Alternative aging phenotypes include disease-free survival, preservation of high levels of function including maintenance of cognitive function (50) and avoidance of bone loss (51), and successful aging (reaching advanced age with intact cognitive ability, physical function, and social engagement) (52). An index of physiologic age developed in the Cardiovascular Health Study by combining data across multiple systems was found to be a better predictor of death and disability than age itself (53). A frailty phenotype defined by five criteria, including unintentional weight loss, exhaustion, weakness, low physical activity, and slow walking speed (54), is distinct from physical disability and itself is predictive of mortality and other adverse outcomes. Although the frailty phenotype was developed in the Cardiovascular Health Study sample, it was found to be applicable across diverse studies (55). Many components of the multidimensional aging phenotypes developed in longitudinal studies are heritable, such as weakness (defined using handgrip strength) and lower extremity function, suggesting the potential for a genetic contribution to the overall phenotype (26,31).

In the family-based Framingham Heart Study, we estimated the heritability of several of the age-related phenotypes including longevity, morbidity-free survival, physical function, and frailty as well as walking speed and handgrip strength (Table 1). For quantitative traits, we used the variance components model, and for dichotomous traits, we used the liability model implemented in the software Sequential Oligogenic Linkage Analysis Routines (56). For both, we defined heritability as the proportion of phenotypic variance due to additive genetic effects only. The heritability of the physical function and frailty phenotypes in the Framingham sample has not previously been reported. For physical disability, three items from the Rosow–Breslau Functional Health Scale (are you able to walk a half mile without help? are you able to walk up and down one flight of stairs without help? are you able to do heavy work around the house without help? [57]) and five items from the Katz Activities of Daily Living Scale (can you do the following five activities independently: dressing, bathing, eating, toileting, and transferring) (58) were used. Physical disability was defined as present if the participant was unable to do any of the items. We examined the presence of physical disability at age 75 years using the exam at which the participant was closest to and within 5 years from age 75 using both the original cohort and offspring samples. Among the 2,614 individuals included in the analysis, 42% reported physical disability at age 75, and the heritability was 44% ($p = .0002$). We estimated heritability of frailty and two of its components handgrip strength and walking speed in the Framingham Offspring cohort participants who attended the last completed research examination (2005–2008) during which the short physical performance battery was administered including a timed 4-m usual paced and quick walk (59). Frailty was defined if three of the five criteria proposed by the Cardiovascular Health Study investigators were present and prefrailty if one to two criteria were present (54). The analysis was adjusted for age and sex and included only participants aged 60 and older. The prevalence of frailty and prefrailty among the 2,207 individuals in this sample was 5% and 41%, respectively, and the combined trait of prefrailty and frailty was modestly heritable ($h^2 = 19\%$, $p = .05$). The usual and fast paced walking times in participants aged 65 years and older were rank normalized to reduce skewness and adjusted for age, sex, body mass index, and height. In contrast to frailty, both the usual and quick walk had a substantial genetic contribution with heritabilities of nearly 40% (usual walk: $h^2 = 0.38$, $p = .0002$; quick walk: $h^2 = 0.36$, $p = .0003$). Next, we estimated heritability of handgrip strength in all offspring participants (mean age 67, range 43–93 years). Handgrip strength was measured three times in each hand with a JAMAR dynamometer. The maximum of the six trials was used in the analysis. Consistent with reports from twin studies, handgrip strength adjusted for age and sex had a heritability of 38% ($p = 5 \times 10^{-15}$). Aging phenotypes are associated with varied heritabilities (Table 1), and thus, the genetic contribution to the phenotype may be quite modest. Populations that differ in environmental factors may produce different heritability estimates even if the genetic factors influencing the trait are the same. Hence, it is remarkable that the heritability estimates for many of the age-related phenotypes are similar. Longevity and age-related phenotypes with higher heritability are of higher priority for genetic association studies, as these phenotypes are more likely to result in multiple genetic associations.

Genetic Association Studies

Genome-wide association studies (GWAS) test genetic variants across the entire genome for association with a phenotype and have proven highly successful for discovery of novel genes and biologic pathways involved in many common complex conditions (Table 2). However, few GWAS of longevity have been conducted to date. The Framingham Heart Study 100K project was the first investigation of the

Table 2. Genetic Association Studies for Human Longevity

Year	Discovery Sample	Replication Sample	Region	Gene	SNP	*p* Value	Odds Ratio
Genome-Wide Association Studies With Discovery and Replication Samples in Humans							
2011 (60)	763 long-lived German individuals (mean age 99.7) 1,085 young German Individuals (mean age 60.2 y)	754 long-lived German individuals (mean age 96.9) 860 young German individuals (mean age 67.3 y)	19q13.32	*APOC1**	rs4420638	1.8×10^{-10}	0.53
2011 (61)	Leiden Longevity Study: 403 Long-lived (mean age 94); 1,760 younger controls (mean age 58)	Rotterdam Study: 960 long-lived (mean age 94); 1,825 younger controls (mean age 62) Leiden 85+ Study: 1,208 long-lived (mean age 92); 2,090 younger controls (mean age 35) Danish 1905 cohort: 1,598 long-lived (mean age 93); 1,997 younger controls (mean age 57)	19q13.32	*TOMM40*†	rs2075650	3.4×10^{-17}	0.71
2011 (62)	410 long-lived individuals from Southern Italy (90–109 y); 553 younger controls (18–48 y)	116 long-lived individuals (90–109 y); 160 younger controls (18–44 y)	5q22.1	*CAMKIV*	rs10491334	1.7×10^{-6}	0.55
2011‖, (63)	CHARGE cohorts (AGES, ARIC, BLSA, CHS, FHS, HABC, InCHIANTI, RS, and SHIP), 25,007 participants age ≥55 y at baseline (55% women), European origin, 8,444 deaths (mean age 81.1); average follow-up 10.6 y	Four independent samples of European origin; *N* = 10,411, deaths = 1,295	3q26.1	*OTOL1*	rs1425609	1.6×10^{-6}	—
2010‖, (64)	CHARGE cohorts (AGES, CHS, FHS, and RS) 1,836 individuals age >90 y; 1,955 individuals who died between ages 55–80 y	Leiden Longevity Study: 940 long-lived (mean age 94); 744 partners of offspring (mean age 60); Danish 1905 cohort: 1,644 long-lived (mean age 93); 2,007 younger Danish twins (mean age 57)	10q23.2	*MINPPI*	rs9664222	6.8×10^{-7}	0.82
Highly Replicated Candidate Gene Association Studies in Humans‡							
2008 (65)	HHP/HAAS: 203 men of Japanese descent who survived to age 95; 402 "average-lived" men who died prior to 81 y	No replication sample *FOXO3A* association subsequently replicated in: German Centenarian study (66); Southern Italian Centenarian study (67); Han Chinese Study in both men and women Southern Chinese Centenarians (68); Danish 1905 cohort of long-lived individuals (69); CHS and Ashkenazi Jewish Centenarians (70)	6q21	*FOXO3a*	rs2802292	.00009	2.75§
1994 (71)	338 French Centenarians; *n* = 160 French adults aged 20–70 y	No replication sample *ApoE E 4* association with longevity subsequently reported in: Danish Centenarians (72); Danish 1905 cohort (73,74)	19q13.32	*Apo E*	*E4 allele*	<.001	0.43

Notes: AGES = Age, Gene/Environment Susceptibility-ReyKjavik Study; ARIC = Atherosclerosis Risk in Communities Study; BLSA = Baltimore Longitudinal Study of Ageing; CHARGE = Cohorts for Heart and Aging Research in Genomic Epidemiology; CHS = Cardiovascular Health Study; FHS = Framingham Heart Study; HAAS = Honolulu Asia Aging Study; HABC = Health, Aging and Body Composition Study; HHP = Honolulu Heart Program; InCHIANTI = Invescchiare nel Chianti; RS = Rotterdam Study; SHIP = Study of Health in Pomerania; SNP = single nucleotide polymorphism.

* Explained by linkage equilibrium with the *ApoE E4* allele (r^2 = .72).

† Explained by moderate linkage disequilibrium with *ApoE E4* (r^2 = .55, rs429358).

‡ For additional candidate genes that may be associated with longevity, see Christensen and coworkers (73) and Barzilai and Gabriely (75).

§ Homozygous minor (GG) versus homozygous major (TT) alleles between cases and controls.

‖ None of the associations achieved genome-wide significance; only the most significant association in the discovery plus replication stage is provided in the table.

GWAS approach for longevity and aging traits (76). The project was relatively small in size including just 1,345 Framingham participants from the largest 310 families and limited in coverage of the genome as the genotyping was conducted with the 100K Affymetrix GeneChip. Modest associations between longevity (defined as age at death) and single nucleotide polymorphisms (SNPs) in or near *FOXO1a*, a gene important for life span in animal models, as well as other candidate genes were observed but failed to reach genome-wide statistical significance. Results of this investigation are considered hypothesis-generating and remain to be replicated. Lending some support to the Framingham 100K longevity investigation, a genome-wide linkage study looking for chromosomal regions linked to successful aging in the Amish Study identified a linkage region near one of the SNPs associated with age at death (52). In 2007, more than 9,300 Framingham Heart Study participants were genotyped with the Affymetrix 500K mapping array plus 50K gene centric supplemental array as part of the National Heart, Lung, and Blood Institute's SNP Health Association Resource project (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000342.v2.p6, accessed March 9, 2012). At the same time, multiple large population-based longitudinal cohort studies in the United States and Europe with richly phenotyped participants planned to conduct genome-wide genotyping. Thus, in 2008, the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium was formed to facilitate GWAS meta-analyses and replication opportunities to enhance gene discovery for many phenotypes (77). The CHARGE aging and longevity working group conducted a meta-analysis of GWAS results for longevity defined as survival to age 90 and older from four cohort studies (Age, Gene/Environment, Susceptibility-Reykjavik Study, the Cardiovascular Health Study, the Framingham Heart Study, and the Rotterdam Study) (64). The CHARGE collaboration permitted the assembly of one of the largest samples of long-lived individuals with genome-wide genotyping available at that time (1,836 individuals achieved aged ≥90 years). The comparison group was drawn from the same cohorts and included only deceased participants to ensure that no individual achieved longevity. The investigation detected 273 SNP associations for longevity that achieved $p < .0001$, but none of the associations achieved genome-wide significance ($p < 5 \times 10^{-8}$). In the next stage of the discovery analysis, among the 24 strongest independent SNP associations in the CHARGE meta-analysis, 16 SNPs were successfully genotyped in the Leiden Longevity Study and the Danish 1905 cohort, and one SNP near the *MINPP1* gene was associated with longevity with $p = 6.8 \times 10^{-7}$ in the combined stage 1 and stage 2 discovery samples. The minor (less frequent) allele was associated with a lower odds of achieving longevity (odds ratio 0.8). *MINPP1* is a highly conserved gene involved in cellular proliferation. The CHARGE aging and longevity working group now includes investigators from over 15 cohort studies permitting future investigations of even larger samples of long-lived individuals with genome-wide genotyping and additional aging phenotypes that may improve our power to detect age-related genetic variation and provide support to our initial findings. We have subsequently conducted a meta-analysis of GWAS data from nine studies in more than 25,000 individuals aged 55 years and older for two age-related phenotypes, all-cause mortality and survival free of major disease and death (63). Although none of the SNP associations for either phenotype achieved genome-wide significance, 14 independent SNPs were associated with mortality, and 8 independent SNPs were associated with event-free survival. The SNPs were in or near genes highly expressed in the brain, genes involved with neural function, and genes associated with a variety of age-related diseases. Thus, our findings suggest that neural processes may be important in regulating aging.

A GWAS conducted in 403 nonagenarians from the Leiden Longevity Study, and 1,670 younger population controls identified 62 SNPs associated with longevity at $p < 1 \times 10^{-4}$ (61). Successful genotyping of 58 of these SNPs was conducted in three independent studies: the Rotterdam Study, Leiden 85-plus Study, and Danish 1905 cohort. A meta-analysis of the 58 SNPs in all four studies that included more than 4,000 nonagenarians and 7,500 younger controls identified only one genome-wide significant SNP rs2075650 in *TOMM40* at chromosome 19q13.32 close to the *ApoE* gene. The minor allele was associated with lower odds of longevity (odds ratio 0.71, $p = 3.4 \times 10^{-17}$). No other SNPs were associated with longevity. SNP rs2075650 was noted to be in moderate linkage disequilibrium with rs429358, the SNP that defines the *ApoE* E4 isoform and in very low linkage disequilibrium with rs7412, the SNP that defines the *ApoE* E2 isoform. In conditional analysis, with all three SNPs in the model, rs2075650 was no longer associated with longevity, whereas the minor allele of rs429358 had a deleterious effect on longevity, and rs7412 a protective effect leading the authors to conclude that rs2075650 effect on longevity was most likely mediated through the isoforms of the *ApoE* gene. A case-control GWAS conducted in 763 German nonagenarians and centenarians, and 1,085 controls (mean age 60 years) identified rs4420638 near the *APOC1* gene and replicated the finding in an independent sample (60). This finding was also fully explained by linkage disquilibrium with *ApoE* E4 isoform confirming the prior report. These results are intriguing as the *ApoE* gene is one of only two candidate genes with consistent evidence for association with longevity in humans (73). The *ApoE* E4 isoform has been linked to elevated cholesterol, cardiovascular disease, age-related cognitive decline, and dementia. The *ApoE* E4 isoform is more strongly associated with Alzheimer disease than longevity and other conditions. Homozygosity for the *Apo E* E4 allele confers up to a 15-fold risk for Alzheimer's disease in whites and an 8-fold risk in African Americans compared with the most

common *ApoE* genotype (E3/E3 [78]). Thus, *ApoE* may influence longevity through premature atherosclerosis and age-related diseases. Notably, the CHARGE study did not observe genome-wide significant associations between the *ApoE* gene region and longevity. However, neither of the two SNPs (rs429358 and rs7412) that define the *ApoE* E4 polymorphism nor any strong proxies appear on any of the chips used by the CHARGE consortium studies. In the CHARGE meta-analysis, the odds of living past age 90 years associated with the minor allele of rs2075650 was 0.85 (p = .046); hence, the effect is consistent with the prior reports.

FOXO3a was first noted to be associated with longevity in a candidate gene study conducted in male centenarians of Japanese descent (65) and subsequently replicated in diverse samples of centenarians and long-lived individuals (66–69). Remarkably neither the Leiden or German studies nor the CHARGE GWAS identified an association between longevity and *FOXO3a*. Finally, a recent GWAS of 410 long-lived individuals and 553 young controls from Southern Italy identified an SNP in an intron of the *CAMKIV* gene among the top associations. This association was replicated in a sample of 116 long-lived and 160 young controls ($p < 10^{-4}$ discovery analysis, joint replication analysis $p = 1.7 \times 10^{-6}$) (62). Interestingly, in vitro work suggests that this gene activates proteins in candidate genes for longevity (*AKT*, *SIRT1*, and *FOXO3a*). About 300 genetic variants in 30 genes in the insulin/insulin-like growth factor 1 (IGF-1) signaling pathway were genotyped in older women participating in the Study of Osteoporotic Fracture. Replication studies were conducted in the Cardiovascular Health Study and an Ashkenazi Jewish Centenarian Study (70). SNPs in two genes in this pathway (*AKT1* and *FOXO3a*) were significantly associated with human life span. A better understanding of the biological mechanisms by which the *FOXO3a* and *AKT1* variants influence human longevity will be important to development of interventions to delay aging (79).

Individuals with a family history of longevity have lower mortality for most age-related diseases (80). Therefore, researchers tested the hypothesis that this may be due to the absence of susceptibility alleles for common diseases in the Leiden Longevity Study and the Leiden 85 Plus Study. They examined whether the long-lived individuals had fewer copies of 30 alleles discovered through GWAS to be associated with coronary artery disease, cancer, and type 2 diabetes compared with a younger comparison group (81). Notably, no difference in the number of risk alleles was detected, suggesting that at least in these populations, survival to old age is not determined by the absence of risk alleles identified to date for age-related disease.

The effort to identify genes that affect longevity through candidate gene and GWAS studies has had only modest success to date. This is likely due to a combination of factors including the heterogeneity of the phenotype, the influence of environmental and dietary factors, which vary widely across populations and the relatively small sample sizes of published longevity GWAS. Many of the successful GWAS with many replicating genome-wide significant signals have sample sizes of more than 10,000, and some published studies of quantitative traits such as age at menarche have had sample sizes of more than 80,000 (43). It is clear that defining more homogeneous phenotypes through age-related traits such as age at menopause and bone mineral density leads to greater success in identifying aging-related genes. Another potential explanation for the lack of identification of risk variants for longevity is that the bulk of the genetic effects are due to rare variants or structural variation in the genome. The GWAS chips used to date have focused on common SNP variants, which typically do not tag rare variants well. Recent work within the CHARGE consortium studies suggests that copy number variants are associated with mortality (82). With the advent of low-cost exome and whole genome sequencing as well as higher-density SNP chips with 5 million or more variants, we will soon have the opportunity to determine whether rare or structural variants explain a substantive proportion of the heritability of longevity and other aging traits.

Translation

Comparison across genomes of different organisms may greatly facilitate the process of gene identification and testing of homologous genes in humans (83). Genome-wide mapping studies of longevity have been conducted in animal models, including *Caenorhabditis elegans* and Drosophila (84). Through the National Institute of Aging's Longevity Consortium (http://longevityconsortium.org/, accessed August 15, 2011) that brings together scientists from multiple disciplines, investigators from the CHARGE consortium, and the Aging Center at the Jackson Laboratory in Maine began to examine the conserved genetic mechanisms between humans and the mouse for life span (85). The laboratory mouse is an excellent model organism for understanding mammalian physiology and genetics for several reasons, including the availability of extensive genetic resources, such as hundreds of inbred, congenic, consomic, and recombinant inbred strains, as well as the ability to add and delete genes via transgenesis and targeted mutagenesis (86,87). Sequence analyses have demonstrated that mice and humans share more than 99% of their genes and that these genes are arranged in a homologous fashion on chromosomes, a phenomenon termed synteny. The appearance of quantitative trait loci (QTL) for a given phenotype in syntenic regions of two different species is evidence that the same gene regulates the phenotype in both species (88). In the past 30 years, seven mouse life-span QTL studies have been carried out in two groups of recombinant inbred strains and 3 four-strain crosses (85).

Although none of the SNP associations for longevity in the CHARGE GWAS achieved genome-wide significance (64), 8 of the top 10 hits were located within a mouse QTL, and five of the human signals were located within 10 Mb of

a mouse QTL peak; the probability that this is due to chance is very low (p = .0025 using Fisher's exact test, based on life-span QTL covering 860 Mb of the 2,700 Mb genome and each human peak being 1 Mb in size) (85). The colocalizations of the human peaks and mouse longevity QTL are especially notable in the distal region of mouse chromosome 1 (summarized in Figure 1). In this region, Gelman (90) first reported a longevity QTL that was identified in recombinant inbred lines of C57BL/6J X DBA/2. Recently, Yuan and coworkers identified two longevity QTL in this region in a backcross mouse population (POHN/DehJ × C57BL/6J) × POHN/DehJ (unpublished data, 2011). The peak of Gelman's QTL overlaps with the peak of one of the Yuan and coworkers QTL, around 160 Mb–163 Mb. The peak of the other Yuan and coworkers QTL was 14 Mb apart, around 175 Mb. Syntenic regions of 2 of the 10 highest peaks (rs16850255 and rs4443878) identified in the CHARGE study are found around 161 Mb and 176 Mb, colocalizing with the mouse longevity QTL. This suggests that human and mouse may share some mechanisms that regulate life span.

Interestingly, around 160 Mb and 175 Mb, one peak from a mouse genome–wide association study and one QTL of IGF-1 have been reported (Figure 1). Yuan and coworkers reported that across mouse inbred strains, lower IGF-1 levels are associated with longer life span. The overlap between longevity QTL and IGF-1 QTL suggests that this region may contain genes that could regulate longevity through the regulation of IGF-1 level. Although this hypothesis must be further verified, the combination of such human and mouse genetic studies establishes a foundation for a powerful translational strategy. We plan to integrate data for additional age-related traits and use bioinformatics and genetic resources in the mouse to test promising candidate genes for longevity. Although genetic association studies in humans can help identify potential genes linked to longevity, the mouse model may be very useful in uncovering the underlying biologic mechanisms that lead to aging.

Other investigators have examined human–chimpanzee orthologous gene pairs to explore evolutionary forces on genes related to aging (93). Genes that appeared to have a pattern of selection tended to have important biological functions that were conserved among mammals (93). Of interest, the study findings suggest that one gene that may have undergone rapid evolution is *WRN*. Defects in *WRN* cause Werner's syndrome characterized by premature aging.

Ultimately, we hope that the identification of genes influencing human longevity will provide insights into the biology of aging and in turn new therapeutics to slow aging and improve health. Evidence from both human and animal studies has identified genes in the IGF1/insulin signaling pathway influencing life span. Therefore, interest in resveratrol, a compound shown to extend life span in animals and improve insulin secretion and insulin sensitivity among many other benefits, is not surprising. A recent pilot study

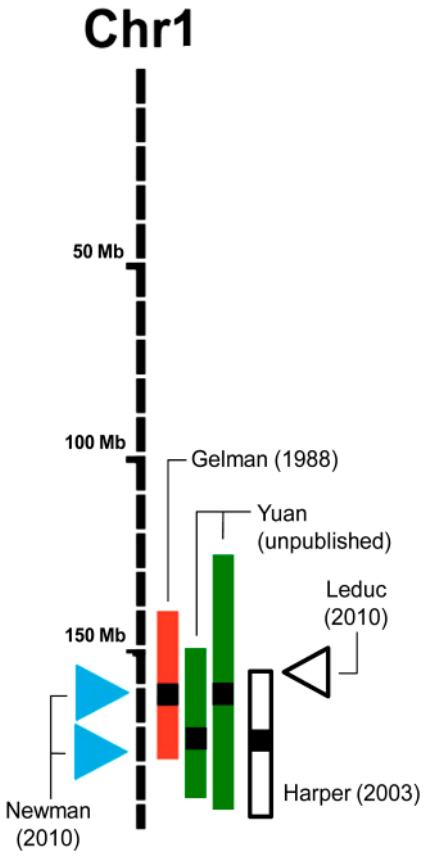


Figure 1. Quantitative trait loci (QTL) for mouse longevity and insulin-like growth factor 1 (IGF-1) as well as genome-wide association peaks for human longevity and mouse IGF-1, both depicted on the mouse genome, mouse chromosome 1 (mapped in Mb). Colored bars are longevity QTL; the open bar is an IGF-1 QTL. The height of the bars represents the 95% confidence interval if reported or an estimated 40 Mb if not reported; the black squares in the bars represent the QTL peaks. We determined the Mb position using a recently revised mouse map (89) and the Mouse Map Converter from the Center for Genome Dynamics (http://cgd.jax.org/mousemapconverter/). Arrows to the left of the chromosome represent human genome-wide association peaks at the homologous mouse genome locations. The arrow on the right of the chromosome is the mouse genome-wide association peak of IGF-1. (Figure is modified from Figure 1 of Yuan et al [85]). Chr = chromosome. Gelman et al (90), Harper et al (91), Leduc et al (92), Newman et al (64).

conducted in 10 older adults suggests that resveratrol improved insulin sensitivity in individuals with impaired glucose tolerance (94). These data combined with animal studies provide support for larger studies of the benefits of resveratrol in humans (95).

Future Directions

Technological advances now permit the complete sequencing of all protein-coding portions of the genome (the "exome"), and the ability to sequence the whole genome of individuals quickly and efficiently is just beginning. Sequencing represents an opportunity to detect rare potentially functional genetic variants unlikely to be discovered with the GWAS that focus on common genetic variation (minor allele frequencies of >5%). The National Human Genome Research Institute and National Heart, Lung, and

Blood Institute (NHLBI) funded the Exome project with the goal of identifying genes contributing to heart, lung, and blood disorders (http://www.nhlbi.nih.gov/resources/exome.htm, accessed August 15, 2011). This innovative technology and the analytic tools under development will be able to be extended to longevity and age-related traits. The sequencing of centenarian genomes may uncover rare genetic variants underlying human extreme longevity and provide insights into the basic biology of aging.

Changes in gene expression that occur as a result of molecular mechanisms that do not change the primary DNA sequence are referred to as epigenetics (http://www.ncbi.nlm.nih.gov/books/NBK45788/#epi_sci_bkgrd.About_Epigenetics, accessed March 9, 2012). Epigenetic mechanisms influence phenotypic expression and are affected by development, the environment, diet, drugs, and aging. One of the best studied epigenetic mechanisms called DNA methylation usually results in suppression of nearby genes. A study of global DNA methylation in an Icelandic cohort and a family-based Utah cohort demonstrated familial clustering of DNA methylation changes and changes in DNA methylation over time (96). The changes in DNA methylation that occur with aging may alter normal gene expression and in turn contribute to development of age-related disease and functional decline (97). Epigenetic changes can be identified using genome-wide analysis with microarrays (ChIP-chip) or next generation sequencing (ChIP-Seq). These new technologies may be used in longitudinal cohort studies in the future to uncover the role of epigenetics in human longevity.

Conclusions

Genetic factors undoubtedly contribute to human aging and longevity, yet candidate gene and GWAS have yielded few replicated longevity-gene associations to date with the exceptions of the *ApoE* and *FOXO3A* genes. Genome-wide genotyping of participants in longitudinal cohort studies, family-based studies, and special populations of long-lived individuals such as centenarians along with unprecedented collaboration among investigators in the United States, Europe, and worldwide provide the opportunity for the assembly of the large discovery and replication samples needed for genetic discoveries. Existing consortia that include both population-based and laboratory-based scientists may speed the translation of newly discovered genetic associations by uncovering the function of the identified genetic variants and ultimately the biologic mechanisms leading to human aging. Many epidemiological studies are poised to use new technologies to move beyond common genetic variants to explore the contribution of low frequency and rare genetic variants, structural changes, and epigenetic changes to human longevity and aging.

Funding

This work was funded through grants from the National Institute on Aging (R01AG29451, Drs J.M.M. and K.L.L.; P30 AG038070, Drs J.M.M. and R.Y.; AG034349, Dr R.Y.).

References

1. Terry DF, Pencina MJ, Vasan RS, et al. Cardiovascular risk factors predictive for survival and morbidity-free survival in the oldest-old Framingham Heart Study participants. *J Am Geriatr Soc*. 2005;53(11):1944–1950.
2. Willcox BJ, He Q, Chen R, et al. Midlife risk factors and healthy survival in men. *JAMA*. 2006;296(19):2343–2350.
3. Dutta A, Henley W, Lang I, et al. Predictors of extraordinary survival in the Iowa established populations for epidemiologic study of the elderly: cohort follow-up to "extinction." *J Am Geriatr Soc*. 2011;59(6):963–971.
4. Yates LB, Djousse L, Kurth T, Buring JE, Gaziano JM. Exceptional longevity in men: modifiable factors associated with survival and function to age 90 years. *Arch Intern Med*. 2008;168(3):284–290.
5. Perls T, Shea-Drinkwater M, Bowen-Flynn J, et al. Exceptional familial clustering for extreme longevity in humans. *J Am Geriatr Soc*. 2000;48(11):1483–1485.
6. Perls TT, Wilmoth J, Levenson R, et al. Life-long sustained mortality advantage of siblings of centenarians. *Proc Natl Acad Sci U S A*. 2002;99(12):8442–8447.
7. Willcox BJ, Willcox DC, He Q, Curb JD, Suzuki M. Siblings of Okinawan centenarians share lifelong mortality advantages. *J Gerontol A Biol Sci Med Sci*. 2006;61(4):345–354.
8. Atzmon G, Schechter C, Greiner W, Davidson D, Rennert G, Barzilai N. Clinical phenotype of families with longevity. *J Am Geriatr Soc*. 2004;52(2):274–277.
9. Terry DF, Wilcox MA, McCormick MA, et al. Lower all-cause, cardiovascular, and cancer mortality in centenarians' offspring. *J Am Geriatr Soc*. 2004;52(12):2074–2076.
10. Newman AB, Glynn NW, Taylor CA, et al. Health and function of participants in the Long Life Family Study: A comparison with other cohorts. *Aging (Albany NY)*. 2011;3(1):63–76.
11. Rozing MP, Westendorp RG, de Craen AJ, et al. Favorable glucose tolerance and lower prevalence of metabolic syndrome in offspring without diabetes mellitus of nonagenarian siblings: the Leiden longevity study. *J Am Geriatr Soc*. 2010;58(3):564–569.
12. Terry DF, Evans JC, Pencina MJ, et al. Characteristics of Framingham offspring participants with long-lived parents. *Arch Intern Med*. 2007;167(5):438–444.
13. Evert J, Lawler E, Bogan H, Perls T. Morbidity profiles of centenarians: survivors, delayers, and escapers. *J Gerontol A Biol Sci Med Sci*. 2003;58(3):232–237.
14. Franceschi C, Motta L, Valensin S, et al. Do men and women follow different trajectories to reach extreme longevity? Italian Multicenter Study on Centenarians (IMUSCE). *Aging (Milano)*. 2000;12(2):77–84.
15. Andersen SL, Sebastiani P, Dworkis DA, Feldman L, Perls TT. Health span approximates life span among many supercentenarians: compression of morbidity at the approximate limit of life span. *J Gerontol A Biol Sci Med Sci*. 2012; doi:10.1093/gerona/glr223.
16. McGue M, Vaupel JW, Holm N, Harvald B. Longevity is moderately heritable in a sample of Danish twins born 1870–1880. *J Gerontol*. 1993;48(6):B237–B244.
17. Herskind AM, McGue M, Holm NV, Sorensen TI, Harvald B, Vaupel JW. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Hum Genet*. 1996;97(3):319–323.
18. Mitchell BD, Hsueh WC, King TM, et al. Heritability of life span in the Old Order Amish. *Am J Med Genet*. 2001;102(4):346–352.
19. Kerber RA, O'Brien E, Smith KR, Cawthon RM. Familial excess longevity in Utah genealogies. *J Gerontol A Biol Sci*. 2001;56(3):B130–B139.
20. Murabito JM, Lunetta KL. Genetics of human longevity and healthy aging. In: Newman AB, Cauley JA, eds. *The Epidemiology of Aging*. New York, NY: Springer; 2011. In press.

21. Lee JH, Flaquer A, Costa R, et al. Genetic influences on life span and survival among elderly African-Americans, Caribbean Hispanics, and Caucasians. *Am J Med Genet A*. 2004;128(2):159–164.
22. Reed T, Dick DM. Heritability and validity of healthy physical aging (wellness) in elderly male twins. *Twin Res*. 2003;6(3):227–234.
23. Christensen K, McGue M, Yashin A, Iachine I, Holm NV, Vaupel JW. Genetic and environmental influences on functional abilities in Danish twins aged 75 years and older. *J Gerontol A Med Sci*. 2000; 55(8):M446–M452.
24. Christensen K, Gaist D, Vaupel JW, McGue M. Genetic contribution to rate of change in functional abilities among Danish twins aged 75 years or more. *Am J Epidemiol*. 2002;155(2):132–139.
25. Dato S, Montesanto A, Lagani V, Jeune B, Christensen K, Passarino G. Frailty phenotypes in the elderly based on cluster analysis: a longitudinal study of two Danish cohorts. Evidence for a genetic influence on frailty. *Age (Dordr)*. 2011.
26. Carmelli D, Reed T. Stability and change in genetic and environmental influences on hand-grip strength in older male twins. *J Appl Physiol*. 2000;89(5):1879–1883.
27. Reed T, Fabsitz RR, Selby JV, Carmelli D. Genetic influences and grip strength norms in the NHLBI twin study males aged 59–69. *Ann Hum Biol*. 1991;18(5):425–432.
28. Frederiksen H, Gaist D, Petersen HC, et al. Hand grip strength: a phenotype suitable for identifying genetic variants affecting mid- and late-life physical functioning. *Genet Epidemiol*. 2002;23(2): 110–122.
29. Matteini AM, Fallin MD, Kammerer CM, et al. Heritability estimates of endophenotypes of long and health life: the Long Life Family Study. *J Gerontol A Biol Sci Med Sci*. 2010;65(12):1375–1379.
30. Ortega-Alonso A, Pedersen NL, Kujala UM, et al. A twin study on the heritability of walking ability among older women. *J Gerontol A Biol Sci Med Sci*. 2006;61(10):1082–1085.
31. Carmelli D, Kelly-Hayes M, Wolf PA, et al. The contribution of genetic influences to measures of lower-extremity function in older male twins. *J Gerontol A Biol Sci*. 2000;55(1):B49–B53.
32. Karasik D, Cupples LA, Hannan MT, Kiel DP. Age, gender, and body mass effects on quantitative trait loci for bone mineral density: the Framingham Study. *Bone*. 2003;33(3):308–316.
33. Karasik D, Myers RH, Cupples LA, et al. Genome screen for quantitative trait loci contributing to normal variation in bone mineral density: the Framingham Study. *J Bone Miner Res*. 2002;17(9): 1718–1727.
34. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry*. 2006; 63(2):168–174.
35. Ashford JW, Mortimer JA. Non-familial Alzheimer's disease is mainly due to genetic factors. *J Alzheimers Dis*. 2002;4(3):169–177.
36. Anderson CA, Duffy DL, Martin NG, Visscher PM. Estimation of variance components for age at menarche in twin families. *Behav Genet*. 2007;37(5):668–677.
37. Towne B, Czerwinski SA, Demerath EW, Blangero J, Roche AF, Siervogel RM. Heritability of age at menarche in girls from the Fels Longitudinal Study. *Am J Phys Anthropol*. 2005;128(1):210–219.
38. Murabito JM, Yang Q, Fox C, Wilson PW, Cupples LA. Heritability of age at natural menopause in the Framingham Heart Study. *J Clin Endocrinol Metab*. 2005;90(6):3427–3430.
39. van Asselt KM, Kok HS, Pearson PL, et al. Heritability of menopausal age in mothers and daughters. *Fertil Steril*. 2004;82(5):1348–1351.
40. de Bruin JP, Bovenhuis H, van Noord PA, et al. The role of genetic factors in age at natural menopause. *Hum Reprod*. 2001;16(9): 2014–2018.
41. Snieder H, MacGregor AJ, Spector TD. Genes control the cessation of a woman's reproductive life: a twin study of hysterectomy and age at menopause. *J Clin Endocrinol Metab*. 1998;83(6):1875–1880.
42. Hjelmborg JV, Iachine I, Skytthe A, et al. Genetic influence on human lifespan and longevity. *Hum Genet*. 2006;119(3):312–321.
43. Elks CE, Perry JR, Sulem P, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet*. 2010;42(12):1077–1085.
44. He C, Kraft P, Chen C, et al. Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat Genet*. 2009;41(6):724–728.
45. Stolk L, Zhai G, van Meurs JB, et al. Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat Genet*. 2009;41(6):645–647.
46. Seshadri S, Fitzpatrick AL, Ikram MA, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA*. 2010;303(18): 1832–1840.
47. Coon KD, Myers AJ, Craig DW, et al. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J Clin Psychiatry*. 2007; 68(4):613–618.
48. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/ MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet*. 2011;43(5):436–441.
49. Rivadeneira F, Styrkarsdottir U, Estrada K, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet*. 2009;41(11):1199–1206.
50. Yaffe K, Lindquist K, Vittinghoff E, et al. The effect of maintaining cognition on risk of disability and death. *J Am Geriatr Soc*. 2010;58(5): 889–894.
51. Cauley JA, Lui LY, Barnes D, et al. Successful skeletal aging: a marker of low fracture risk and longevity. The Study of Osteoporotic Fractures (SOF). *J Bone Miner Res*. 2009;24(1):134–143.
52. Edwards DR, Gilbert JR, Jiang L, et al. Successful aging shows linkage to chromosomes 6, 7, and 14 in the Amish. *Ann Hum Genet*. 2011; 75(4):516–528.
53. Newman AB, Boudreau RM, Naydeck BL, Fried LF, Harris TB. A physiologic index of comorbidity: relationship to mortality and disability. *J Gerontol A Biol Sci Med Sci*. 2008;63(6):603–609.
54. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Med Sci*. 2001;56(3):M146–M156.
55. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the women's health and aging studies. *J Gerontol A Biol Sci Med Sci*. 2006;61(3):262–266.
56. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet*. 1998;62(5):1198–1211.
57. Rosow I, Breslau NA. Guttman health scale for the aged. *J Gerontol*. 1966;21(4):556–559.
58. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc*. 1983; 31(12):721–727.
59. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*. 1994;49(2):M85–M94.
60. Nebel A, Kleindorp R, Caliebe A, et al. A genome-wide association study confirms APOE as the major gene influencing survival in long-lived individuals. *Mech Ageing Dev*. 2011;132(6–7):324–330.
61. Deelen J, Beekman M, Uh HW, et al. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. *Aging Cell*. 2011;10(4):686–698.
62. Malovini A, Illario M, Iaccarino G, et al. Association study on long-living individuals from Southern Italy identifies rs10491334 in the CAMKIV gene that regulates survival proteins. *Rejuvenation Res*. 2011;14(3):283–291.
63. Walter S, Atzmon G, Demerath EW, et al. A genome-wide association study of ageing. *Neurobiol Aging*. 2011;32(11):2109e15–28.
64. Newman AB, Walter S, Lunetta KL, et al. A meta-analysis of four genome-wide association studies of survival to age 90 years or older: the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. *J Gerontol A Biol Sci Med Sci*. 2010;65(5): 478–487.

65. Willcox BJ, Donlon TA, He Q, et al. FOXO3A genotype is strongly associated with human longevity. *Proc Natl Acad Sci U S A*. 2008; 105(37):13987–13992.
66. Flachsbart F, Caliebe A, Kleindorp R, et al. Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc Natl Acad Sci U S A*. 2009;106(8):2700–2705.
67. Anselmi CV, Malovini A, Roncarati R, et al. Association of the FOXO3A locus with extreme longevity in a southern Italian centenarian study. *Rejuvenation Res*. 2009;12(2):95–104.
68. Li Y, Wang WJ, Cao H, et al. Genetic association of FOXO1A and FOXO3A with longevity trait in Han Chinese populations. *Hum Mol Genet*. 2009;18(24):4897–4904.
69. Soerensen M, Dato S, Christensen K, et al. Replication of an association of variation in the FOXO3A gene with human longevity using both case-control and longitudinal data. *Aging Cell*. 2010;9(6):1010–1017.
70. Pawlikowska L, Hu D, Huntsman S, et al. Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell*. 2009;8(4):460–472.
71. Schachter F, Faure-Delanef L, Guenot F, et al. Genetic associations with human longevity at the APOE and ACE loci. *Nat Genet*. 1994; 6(1):29–32.
72. Gerdes LU, Jeune B, Ranberg KA, Nybo H, Vaupel JW. Estimation of apolipoprotein E genotype-specific relative mortality risks from the distribution of genotypes in centenarians and middle-aged men: apolipoprotein E gene is a "frailty gene," not a "longevity gene." *Genet Epidemiol*. 2000;19(3):202–210.
73. Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet*. 2006;7(6):436–448.
74. Bathum L, Christiansen L, Jeune B, Vaupel J, McGue M, Christensen K. Apolipoprotein e genotypes: relationship to cognitive functioning, cognitive decline, and survival in nonagenarians. *J Am Geriatr Soc*. 2006;54(4):654–658.
75. Barzilai N, Gabriely I. Genetic studies reveal the role of the endocrine and metabolic systems in aging. *J Clin Endocrinol Metab*. 2010; 95(10):4493–4500.
76. Lunetta KL, D'Agostino RB Sr, Karasik D, et al. Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. *BMC Med Genet*. 2007; 8(suppl 1):S13.
77. Psaty BM, O'Donnell CJ, Gudnason V, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2(1):73–80.
78. Logue MW, Schu M, Vardarajan BN, et al. A comprehensive genetic association study of Alzheimer disease in African Americans. *Arch Neurol*. 2011;68(12):1569–1579.
79. Tazearslan C, Cho M, Suh Y. Discovery of functional gene variants associated with human longevity: opportunities and challenges. *J Gerontol A Biol Sci Med Sci*. 2011; doi:10.1093/gerona/glr200.
80. O'Brien E, Kerber R, Smith K, Mineau G, Boucher K, Reed DL. Familial mortality in the Utah population database: characterizing a human aging phenotype. *J Gerontol A Biol Sci Med Sci*. 2007;62(8): 803–812.
81. Beekman M, Nederstigt C, Suchiman HE, et al. Genome-wide association study (GWAS)-identified disease risk alleles do not compromise human longevity. *Proc Natl Acad Sci U S A*. 2010;107(42): 18046–18049.
82. Kuningas M, Estrada K, Hsu YH, et al. Large common deletions associate with mortality at old age. *Hum Mol Genet*. In press.
83. Burgess-Herbert SL, Cox A, Tsaih SW, Paigen B. Practical applications of the bioinformatics toolbox for narrowing quantitative trait loci. *Genetics*. 2008;180(4):2227–2235.
84. Shmookler Reis RJ, Kang P, Ayyadevara S. Quantitative trait loci define genes and pathways underlying genetic variation in longevity. *Exp Gerontol*. 2006;41(10):1046–1054.
85. Yuan R, Peters LL, Paigen B. Mice as a mammalian model for research on the genetics of aging. *ILAR J*. 2011;52(1):4–15.
86. Paigen K, Eppig JT. A mouse phenome project. *Mamm Genome*. 2000;11(9):715–717.
87. O'Brien T, Woychik R. Our small relative. *Nat Genet*. 2003;1:3–4.
88. Sugiyama F, Churchill GA, Higgins DC, et al. Concordance of murine quantitative trait loci for salt-induced hypertension with rat and human loci. *Genomics*. 2001;1:70–77.
89. Cox A, Ackert-Bicknell CL, Dumont BL, et al. A new standard genetic map for the laboratory mouse. *Genetics*. 2009;182(4): 1335–1344.
90. Gelman R, Watson A, Bronson R, Yunis E. Murine chromosomal regions correlated with longevity. *Genetics*. 1988;118(4):693–704.
91. Harper JM, Galecki AT, Burke DT, Pinkosky SL, Miller RA. Quantitative trait loci for insulin-like growth factor I, leptin, thyroxine, and corticosterone in genetically heterogeneous mice. *Physiol Genomics*. 2003;15(1):44–51.
92. Leduc MS, Hageman RS, Meng Q, et al. Identification of genetic determinants of IGF-1 levels and longevity among mouse inbred strains. *Aging Cell*. 2010;9(5):823–836.
93. de Magalhaes JP, Church GM. Analyses of human-chimpanzee orthologous gene pairs to explore evolutionary hypotheses of aging. *Mech Ageing Dev*. 2007;128(5–6):355–364.
94. Crandall JP, Oram V, Trandafirescu G, et al. Pilot study of resveratrol in older adults with impaired glucose tolerance. *J Gerontol A Biol Sci Med Sci*. 2012; doi:10.1093/gerona/glr235.
95. Smoliga JM, Vang O, Baur JA. Challenges of translating basic research into therapeutics: resveratrol as an example. *J Gerontol A Biol Sci Med Sci*. 2011;67(2):158–167.
96. Bjornsson HT, Sigurdsson MI, Fallin MD, et al. Intra-individual change over time in DNA methylation with familial clustering. *JAMA*. 2008;299(24):2877–2883.
97. Gravina S, Vijg J. Epigenetic factors in aging and longevity. *Pflugers Arch*. 2010;459(2):247–258.

CLINICAL MANAGEMENT extra

Comprehensive Programs for Preventing Pressure Ulcers: A Review of the Literature

CME
CATEGORY 1
1 Credit

CE
ANCC
2.9 Contact Hours

Andrea Niederhauser, MPH • Program Manager • Boston University School of Public Health • Boston, Massachusetts
Carol VanDeusen Lukas, EdD • Senior Investigator • Center for Organization, Leadership, and Management Research • Veterans Health Administration • Boston Healthcare System • Clinical Associate Professor • Boston University School of Public Health • Boston, Massachusetts
Victoria Parker, DBA • Associate Professor • Health Policy and Management • Boston University School of Public Health • Boston, Massachusetts
Elizabeth A. Ayello, PhD, RN, ACNS-BC, CWON, MAPWCA • FAAN • Faculty • Excelsior College of Nursing • Albany, New York • Senior Advisor • The John A. Hartford Institute for Geriatric Nursing • New York, New York • President • Ayello, Harris & Associates • New York, New York • Clinical Editor • *Advances in Skin & Wound Care* • Ambler, Pennsylvania
Karen Zulkowski, DNS, RN •Associate Professor • Montana State University-Bozeman • Bozeman, Montana
Dan Berlowitz, MD, MPH • Director • Center for Health Quality, Outcomes and Economic Research • Bedford Veterans Affairs Hospital • Professor • Boston University School of Public Health • Boston, Massachusetts

All staff, authors, and planners, including spouses/partners (if any), in any position to control the content of this CME activity have disclosed that they have no financial relationships with, or financial interests in, any commercial companies pertaining to this educational activity.

To earn CME credit, you must read the CME article and complete the quiz and evaluation on the enclosed answer form, answering at least 12 of the 17 questions correctly.

This continuing educational activity will expire for physicians on April 30, 2013.

PURPOSE:

To enhance the learner's competence in pressure ulcer (PrU) prevention through a literature review of comprehensive programs.

TARGET AUDIENCE:

This continuing education activity is intended for physicians and nurses with an interest in skin and wound care.

OBJECTIVES:

After participating in this educational activity, the participant should be better able to:

1. Analyze the findings of the PrU prevention program studies found in the literature review.

2. Apply research findings to clinical practice.

ABSTRACT

OBJECTIVE: The objective of this study was to examine the evidence supporting the combined use of interventions to prevent pressure ulcers (PrUs) in acute care and long-term-care facilities.

DESIGN: A systematic review of the literature describing multifaceted PrU prevention programs was performed. Articles were included if they described an intervention implemented in acute care settings or long-term-care facilities, incorporated more than 1 intervention component, involved a multidisciplinary team, and included information about outcomes related to the intervention.

MAIN RESULTS: Twenty-four studies were identified. Recurring components used in the development and implementation of PrU prevention programs included preparations prior to the start of a program, PrU prevention best practices, staff education, clinical monitoring and feedback, skin care champions, and cues to action. Ten studies reported PrU prevalence rates; 9 of them reported decreased prevalence rates at the end of their programs. Of the 6 studies reporting PrU incidence rates, 5 reported a decrease in incidence rates. Four studies measured care processes: 1 study reported an overall improvement; 2 studies reported improvement on some, but not all, measures; and 1 study reported no change.

CONCLUSIONS: There is a growing literature describing multipronged, multidisciplinary interventions to prevent PrUs in acute care settings and long-term-care facilities. Outcomes reported in these studies suggest that such programs can be successful in reducing PrU prevalence or incidence rates. However, to strengthen the level of evidence, sites should be encouraged to rigorously evaluate their programs and to publish their results.

KEYWORDS: preventing pressure ulcers, multidisciplinary interventions to prevent pressure ulcers, reducing incidence of pressure ulcers

ADV SKIN WOUND CARE 2012;25:167-88; quiz 189-90.

SCOPE OF THE PROBLEM

Pressure ulcers (PrUs) remain a national priority in American healthcare. Every year, tens of thousands of patients develop these skin lesions. Although rates vary widely among different care settings, a 2009 survey among self-selected sites in the United States found an overall PrU prevalence rate of 11.9% and a facility-acquired rate of 5.0% in acute care facilities.[1] Along with pain and the risk for serious infections, PrUs result in increased healthcare utilization and costs. A survey using Medicare inpatient data found that, between 2005 and 2007, PrUs accounted for up to $2.41 billion in excess healthcare costs.[2]

Some PrUs may be avoidable, and in the past decade, the prevention of PrUs has gained increased emphasis in clinical practice. In the quest to reduce harm to patients from serious preventable events, institutions such as the National Quality Forum, the Agency for Healthcare Quality and Research, and the Joint Commission, among others, selected PrUs as indicators for patient safety and quality of hospital care. In similar efforts, the 5 Million Lives Campaign, led by the Institute for Healthcare Improvement, brought additional attention to the importance of preventing PrUs in healthcare facilities. Furthermore, in 2008, the Centers for Medicare & Medicaid Services implemented new payment rules that included PrUs on the list of ''never events'' and stopped reimbursement to hospitals for the costs of care resulting from facility-acquired Stages III and IV PrUs.[3]

This article will help clinicians identify best practice evidence supporting the combined use of interventions to prevent PrUs in acute care and long-term-care facilities.

PRESSURE ULCER BEST PRACTICES

In the United States, best practices to prevent PrUs have been identified in randomized controlled trials and have been widely disseminated through clinical practice guidelines. A number of systematic studies have evaluated the efficiency of individual best practices, such as the use of special support surfaces or standardized tools to assess risk for PrU development.[4,5] However, the systematic implementation of best practices one at a time in the standard care environment has been shown to be a challenge for many facilities. Studies have found low adherence rates to best practices for PrU prevention in different care settings. For example, 1 study found that although hospitalized older adults were assessed for PrU risks, only 15% had a supportive device in place by day 3 of hospitalization.[6] Another study found that, of 2425 hospitalized Medicare beneficiaries from across the nation, only 23% of immobile patients were documented as being at risk within 48 hours of admission, 66% were repositioned every 2 hours, and 8% received a pressure-reducing device.[7] A third study surveyed medical records of 834 residents in Veterans Health Administration long-term-care facilities and found that overall adherence to 6 critical best-care practices (such as standardized risk assessment and regular repositioning) was only 50%.[8] Rather than implementing single best practices, care facilities have increasingly begun to bundle best practices together and implement them as part of comprehensive programs in their facilities. The aim of this review was to examine the evidence supporting the combined use of multiple interventions to prevent PrUs in acute care and long-term-care facilities. Two questions guided the analysis of the literature:

- Are there any specific components that have consistently been included in multifaceted programs?
- Is there evidence that these comprehensive programs reduce PrU incidence and/or prevalence?

METHODS

For this review, Ovid MEDLINE and Ovid CINAHL were searched using combinations of the following search terms: *pressure ulcer*, *bed sore*, *decubitus ulcer*, *prevention*, *protocol*, *best practice*, *quality assurance*, and *tool*. In addition, reference lists were reviewed, and clinical experts were contacted to identify further relevant studies. The search was limited to articles published in English between January 1995 and December 2010. To be included, the studies needed to describe a program to prevent PrUs that

- was implemented in an acute care or long-term-care facility,
- consisted of more than 1 intervention component,
- was not limited to site-specific PrUs (such as heel ulcers),
- was delivered through multidisciplinary efforts, and
- measured and reported PrU prevalence or incidence rates before and after implementation of the program.

RESULTS

Twenty-four articles were identified describing comprehensive PrU prevention programs. Twenty studies described single-site interventions and four described multisite interventions. All of the reviewed studies used a longitudinal 1-group pretest-posttest design. No randomized controlled trials were reported. An in-depth review of the studies was performed, and each study was analyzed for the following elements: setting and scope of the program, implementation team and preparations prior to program implementation, intervention components, methods of data collection, and results. The different intervention components were categorized into the following groups: PrU prevention best practices, staff education, clinical monitoring and evaluation, skin care champions, other campaign elements, and strategies to ensure sustainability. The findings of this review are shown in Table 1 and are summarized in the following sections.

Setting and Scope

Of the 20 programs reviewed in acute care settings,[9–28] all but one[17] were rolled out on multiple units or hospital-wide. Two studies reported spreading the program throughout the system after testing it on a small number of pilot units.[18,22] Long-term-care facility initiatives included between 1 and 20 participating facilities.[29–32]

Team

The individuals who initiated and led the improvement efforts were not specified in any of the reviewed articles. However, a distinction was found between programs that were initiated and led internally by staff members[9–13,15–28] and programs that were designed by external experts and implemented in collaboration with the facility in question.[14,29–32] Once the need for change was identified, several facilities chose to establish a team responsible for the design and/or implementation of the intervention. Teams responsible for designing and implementing the programs were generally multidisciplinary and, if specified, included combinations of nurses and nursing aides, wound experts, dietitians, pharmacists, physical therapists, physicians, clinical researchers, educators, information technology staff, managers, and directors.[10,11,15,18,20–25,27]

Program Components

Preparations. Twelve of the reviewed studies specified a set of activities completed prior to developing and implementing their initiatives.[10–13,15,16,18,20,22,24,27,31] Preparations included literature reviews of best practices for prevention and treatment of PrUs; baseline prevalence and incidence surveys; assessments of current state of staff knowledge, existing policies, and care processes; and the testing, evaluation, and selection of pressure relief equipment or skin care products.

Pressure Ulcer Prevention Best Practices. Consistent with existing clinical practice guidelines, studies that described their PrU prevention protocol most commonly reported the use of standardized tools for assessing risk for PrUs; regular skin (re)assessment; an individualized care plan for patients at risk for PrUs; the use of pressure relief equipment, such as low-air-loss mattresses and heel lifts; nutritional assessment and consultation for at-risk patients; frequent turning and repositioning; and the use of skin care products and moisture barrier creams.

Staff Education. All but 4 studies[13,16,23,30] described some form of education or training to increase staff knowledge. Generally, education was targeted at nursing staff and included instructions on PrU treatment and prevention practices, presentations of new or existing facility guidelines and policies, and training on the use of skin care products and support surfaces. Two studies reported educational programs for physicians.[20,28] Most of the studies reported formal staff educational activities, such as unit in-service sessions and workshops, computerized educational modules, educational packages for staff, skin fairs, and wound conferences.[9–12,14,15,18–22,24–29,31,32] Three studies furthermore reported the integration of PrU prevention and treatment into their orientation of new hires.[18,20,21]

In addition to formal activities, several studies described more informal ways of teaching PrU prevention.[9,11,17,22,27,31] One-on-one mentoring, consultation, and support at the bedside through certified nursing staff, training provided in preparation of PrU data collection, and individual case reviews of hospital-acquired PrUs were among the opportunities seized to provide ongoing training.

Clinical Monitoring and Feedback. Ongoing clinical monitoring was frequently used to encourage behavior change and ensure compliance with the PrU prevention practices. Staff compliance with existing protocols was monitored through daily, weekly, or monthly rounding of the wound nurses or other nursing staff; regular chart audits; and PrU tracking forms and compliance monitoring tools.[9–13,15,18,21,22,24,26–28,30,31] Additional strategies for continuous improvement included preventability or root-cause analyses when a patient developed a PrU and the development of action plans when results of the surveys were unsatisfactory.[13,18,26]

Table 1.
STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER)

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
Acute care setting						
9	Baldelli and Paciella, 2008	Hospital intensive care unit (ICU) medical-surgical (United States)	Task force (not specified)	Not reported	Risk assessment Skin assessment Head of bed <30 degrees Incontinence skin care Turning, positioning Heel elevation Nutritional assessment Pressure relief	Presentation (at launch) Bedside education (ongoing)
10	Bales and Padwojski, 2009	Hospital-wide (United States)	Wound committee: CWOCN wound champions	CWOCN hours increased to full time Analysis of care processes and PrU protocols Literature review on evidence-based practice interventions	Computer tool for PrU assessment and initial care, including skin tear protocol Pressure relief surfaces evaluated and purchased PrU prevention algorithm for surgical patients	Mandatory education sessions
11	Catania et al, 2007	All 5 inpatient units at hospital (Cancer hospital; USA)	Multidisciplinary QI team: Clinical nurse specialist (CNS) Nursing director ET registered nurse (RN) quality manager Staff development specialist	*Initial intervention*: chart audits by CNS *Second intervention*: assessment of state of practice	Risk assessment Skin assessment Individualized care plan based on Braden subscores and linked to products/services available at the facility	*Initial intervention:* Unit in-service sessions *Second intervention:* Information packets (at launch) CNS available to assist and mentor staff

Providing feedback on the quality improvement process and sharing PrU rates with the staff at all levels was a priority reported in the majority of studies. Eight of the 24 studies highlighted the importance of sharing the results from their surveys with the staff by posting the PrU rates on unit billboards; publishing them in unit newsletters; distributing them to unit managers, directors, and senior leadership; or discussing them at staff meetings.[9,12,17,18,20,21,30,32]

Skin Care Champions. Seven of the acute care studies created the role of a skin care champion as part of their PrU prevention

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
Compliance monitoring: daily unit audits Posting of rates on units and discussion during meetings Continuous support from ET team	Not reported	Turn clocks Theme: ''Check, Rock & Roll Around the Clock''	Continued data collection	Annual 2-d prevalence/ incidence study *For project:* monthly prevalence/incidence study Unit-specific action plans if needed 2006: Overall prevalence: 15% Overall incidence: 7% (decreased from year before) Monthly studies: below national benchmark, downward trend	Not reported
Weekly audits on each unit Event forms for each HAPU Daily review of documentation and skin assessments for patients with Braden scores <19	Unit-based wound champions	Theme song	Two campaigns to maintain staff motivation	Quarterly prevalence studies Monthly incidence rates based on event forms for every HAPU *Baseline* August 2007 Overall prevalence: 9.5% *Post* December 2008 Overall prevalence: 5.66% HAPU Prevalence: 0%	Not reported
Second intervention: weekly chart audits RN: completed protocol PCA: audits chart 3 d later	Not reported	Not reported	Continued data collection	Quarterly prevalence studies *Baseline*: May 2003 All: 19.47% HAPUs: 12.39% *Post*: December 2004 All: 4.11% HAPUs: 2.05% Weekly monitoring of at-risk patients and staff compliance with protocol	Not reported

Table 1.

STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
12	Chicano and Drolshagen, 2009	25-bed intermediate-care unit (United States)	Quality council members (not specified)	Assessed staff knowledge Assessed current practices Reviewed existing tools and standards of practice	Risk assessment Skin assessment Repositioning schedule Electronic documentation system Revised practice standards for the use of TED stockings and SCDs Updated wound assessment guidelines	Staff education
13	Courtney et al, 2006	Hospital-wide (United States)	Not reported	Assessment of performance in regard to PrU Assessment of potential gaps/ problems	Risk and skin assessment in OR Revised skin breakdown prevention protocol Support equipment	Not reported
14	De Laat et al, 2006	Hospital-wide (the Nether lands)	Not reported	Not reported	Hospital guideline for PrU care Pressure-reducing mattresses	Introduction of new guidelines during staff meetings or clinical lesson

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
Chart audits Staff adherence to guidelines reported during staff meetings	Not reported	Stop skin sign alerts placed on patient charts	Not reported	Quarterly prevalence studies HAPU Incidence: 2005: 6 occurrences within 12 mo 2006: 5 occurrences within 12 mo 2007: 0 occurrences within 12 mo	Not reported
Prevent ability analysis Chart audits Outcome improvement plans	SOS champion on each unit	PrU pocket guide Theme song Stickers on patient charts Redefinition of roles and responsibilities of process owners	Ongoing monitoring of process measures	Baseline prevalence/ incidence study 2001 Overall prevalence: 13% 2001 HAPU incidence: 9.4% 2005 HAPU incidence: 1.8%	Not reported
Not reported	One contact nurse on every unit	Not reported	Continued data collection	*Baseline* HAPU rate: 18% 4-mo *Post* HAPU rate: 13% 11-mo *Post:* HAPU rate: 11%	*Baseline* Inadequate preventive measures: 19% Inadequate treatment: 60% Adequate repositioning: 7% 4-mo Post inadequate preventive measure: 4% Inadequate treatment: 31% Adequate repos: 10% 11-mo Post: inadequate preventive measures: 6% Inadequate treatment: 31%

Table 1.

STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
15	Dibsie, 2008	Hospital-wide (United States)	Multidisciplinary team (not specified)	Assessed current practices	Standardization of products Organization of supply carts on each unit	In-service education
16	Elliott 2010	Three acute-care sites (United Kingdom)	Not reported	Review of previous and current prevention activities to identify areas for improvement	Risk assessment Positioning and repositioning documentation Focus on heel ulcer prevention Purchased 350 dynamic mattresses for high-risk patients	Not reported
17	Elliott et al, 2008	14-bed adult ICU (Australia)	Not reported	Not reported	Not reported	One-on-one training (at launch)
18	Gibbons et al, 2006	Hospital-wide (United States) Three pilot units Staggered rollout	Leadership team (SKIN team): CNO Nurse manager Educator Pharmacist Dietitian 2 staff RNs 2 WOCN Nurse in performance improvement Long-term-care nursing educator	Assessed current practices Literature review Expert meeting Initial chart review	SKIN bundle Surfaces Keep patients turning Incontinence Nutrition	Initial education (at launch) Ongoing education for continuing and new staff Self-study modules

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
Weekly rounding and documentation wounds	Skin champions representing all areas of the hospital	Newsletter	Not reported	Quarterly prevalence studies *Baseline*: All: 15% HAPUs: 9% *Post:* HAPU in 2007: 3% (stage 2 and higher)	Not reported
Not reported	Tissue viability support workers at each site	Not reported	Project has identified areas in need for continued improvement efforts	Annual prevalence study *Baseline*: February 2009: 15.5% *Post*: February 2010: 13.4%	Documentation of repositioning regimen increased from 6% to 13% Documented evidence for repositioning decreased from 11% to 10% Patients nursed on appropriate mattress increased from 79% to 88% Use of heel protectants increased from 33% to 36%
Presentation of data	Not reported	Reminders (not specified) Monthly newsletter	Not reported	Monthly prevalence studies *Baseline*: 2003 Prevalence: 50% *Post*: 2005 Prevalence: 8%	Not reported
Compliance monitoring tool (not specified) Root-cause analysis Review results at weekly skin operations meetings	Not reported	SKIN risk reminders placed on nursing clipboards Newsletters Poster Pocket reference cards	Continued data collection Identification of other opportunities for PrU prevention	Quarterly prevalence studies Annual incidence study HAPU baseline prevalence: 5.7% Incidence: 2004: >2% 2006: <1% No Stage III and IV PrUs between 2004 and 2006	Not reported

Table 1.

STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
19	Gunningberg and Stotts, 2008	Hospital-wide (Sweden)	Not reported	Not reported	Development of clinical guidelines (not specified) Mandatory use of documentation templates	Education for nurses and nurse assistants Web-based program
20	Hiser et al, 2006	Five participating units (not specified) (United States)	Wound care team: WOCN ARNP Department manager RNs Physical therapists Dietitians Skin resource team	Lit review Assessed policies Assessed knowledge	Care planning linked to Braden score PrU prevention protocol PrU/skin tear physician orders Risk assessment with Braden Scale Dietary consults Ostomy supplies updated Support surfaces evaluated and changed List of skin care products standardized	Annual wound conference Physician education New staff orientation Skills fairs 6 times a year

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
		SKIN reminders on patient chart Weekly SKIN operations meeting Medical staff approved standing order for dietitians			
Not reported	Not reported	Not reported	Not reported	Cross-sectional prevalence studies *2002 Baseline*: 23.9% (all stages) 8% (Stages II-IV) *2006 Post*: 22.9% (all stages) 12% (Stages II-IV) No statistically significant decrease	2006: Retrospective chart audit Skin inspection recorded: 41.2% Risk assessment documented 25%
Timely feedback to staff	Not reported	Redefinition of role of WOCN Skin resource team Newsletter Resource book	Not reported	Quarterly Prevalence-studies 2002 HAPU Prevalence: 9.2% 2004 HAPU Prevalence: 6.6%	Not reported

Table 1.

STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
21	Hopkins et al, 2000	Medical-surgical unit (United States)	PrU team (planned the surveys): nurse researcher, skin care clinical nurse specialists, nurse educator, staff nurse, QI representative	After each annual survey, implementation of new practices	Use of Braden scale AHRQ guidelines for protocols Use of specialty equipment Revise/update protocol	Continuing education activities Nursing orientation PrU program
22	LeMaster, 2007	Pulmonary and oncology units (United States)	CNS Group	Analyzed HAPU data to identify trends and areas for improvement Baseline knowledge assessment	Standard intervention TOE: *T*urn, *O*verlay, *E*levate	Education at regular staff meeting (at launch) Direct consultation services to bedside staff
23	McInerney, 2008	Hospital-wide (United States)	Heel intervention: task force: critical care physician, podiatrist, risk manager, 2 WOCNs, CMO, CNO, clinical informatics analyst, manager of central distribution, education OR and CC	After 18 mo: Review of literature and inspection of heel pressure-relieving devices	*First intervention* Risk assessment Static air overlay for patients with score <16 *After 18 mo:* Heelift boots Powered air beds for critical care Pressure-reducing mattresses	Not reported
24	Sacharok and Drew, 1998	Adult medical, surgical, and critical care units (United States)	Skin care team: RNs, nursing administration, unit manager, WOCN, clinical specialist ''Rear admirals'' (skin care resource person, nursing unit representative)	Literature review Knowledge and care-delivery patterns assessment Retrospective chart review of PrU incidence ''Gap'' analysis	PrU protocol developed and approved for use hospital-wide Pressure-reducing mattresses tested and purchased Skin care products evaluated Braden scale on admission and weekly	Education at staff meetings Mandatory annual equipment fairs Skin care fair

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
Display study results/ present at meetings and conferences Monitor documentation and usage of specialty beds (not specified)	Unit skin resource nurse	Quick reference of protocols	Not reported	HAPU Prevalence 1996: 18% 1997: 10% 1998: 9%	Not reported
Daily rounding for 6 wk	Not reported	TOE acronym placed on top of turn sheets and algorithms Internal resources manual	Implementation of additional cues to help with maintaining practice	Quarterly prevalence studies Prevalence rate decreased to zero after implementation	Not reported
Not reported	Not reported	*First intervention* Second WOCN hired	Continued data collection	Semiannual HAPU prevalence study February 2002: HAPU prevalence: 12.8% July 2003: HAPU prevalence: 5.1% July 2005: HAPU prevalence: 2% July 2006: HAPU prevalence: 2.4%	Not reported
Monthly prevalence rounds by rear admirals	Not reported	Poster on units Pocket reference guide	Continued data collection Identification of other opportunities for PrU prevention	1994: Monthly prevalence studies 1995–1997: Quarterly prevalence studies 1994: HAPU prevalence: 19% 1997: HAPU prevalence: 3% Retrospective chart review 1995: HAPU incidence rate 43% lower than in 1994	Not reported

Table 1.
STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
					Multidisciplinary consultations (nutrition, physical therapy, ET) Nursing care flow sheet redesign Standardization of products	
25	Stausberg et al, 2006 and 2009	Hospital-wide (Germany)	Interdisciplinary team: Nurses Physicians Information technology specialists and researchers	Not reported	PrU prevention guidelines Risk assessment and derivation of preventive measures Optimization of pressure system supply Introduction of special foam mattresses	Staff training
26	Stoelting et al, 2007	Hospital-wide (United States)	Wound ostomy management team (not specified)	Not reported	Braden Scale for risk assessment PrU prevention protocol	Staff education
27	Young et al, 2010	Hospital-wide (United States)	Clinician-led task force	Literature review Assessment of stakeholders and environmental readiness for change	Revised and updated skin care policy New support surfaces evaluated and selected Standardized patient turning schedule New skin care products	Education of nurses through mandatory continuing educational sessions Monthly updates from unit champions
28	Young et al, 2003	Acute care, rehabilitation center, and skilled nursing unit (United States)	2 ET nurses (not specified)	Not reported	Braden scale Weekly/biweekly assessment of patients at risk Standard protocol with interventions	In-service Hospital-wide conferences Physician education

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
Not reported	Not reported	Not reported	Not reported	Retrospective analysis of electronic hospital data Point prevalence rates increased significantly from 1.44% in year 1 to 1.77% in year 4 Incidence rate increased from 0.56% to 0.65%	Not reported
PrU tracking form Individual case reviews and plans of actions	Wound ostomy unit champion	Red Dot initiative (not specified)	Continued data collection	Prevalence study Second inspection of patients 4 d later Incidence rate dropped from 7% to 4%	Not reported
Monthly chart audits	Unit champions	*Save Our Skin* logo Laminated copy of Braden scale placed at each patient's bedside	Not reported	Prevalence and incidence study: Fall 2006: HAPU incidence Campus 1: 12.5% Campus 2: 8.7% Campus 3: 0% Spring 2007: HAPU incidence Campus 1: 9.1% Campus 2: 2.8% Campus 3: 0%	Not reported
Weekly prevention assessment rounds	Not reported	Expanded role of ET nurse from treatment to prevention Increased from 1 to 2 full-time equivalent ET	Identification of other opportunities for PrU prevention	Quarterly prevalence walks (results not reported) Incidence measured 1999–2001: 55% decrease in PrU incidence	Compliance with skin assessment on admission: increased to 97%

Table 1.
STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
					Highlighted in patient chart Individual turn schedules ET nurse PrU consult sheets Evaluation/ monitoring of specialty products	
Long-term-care setting						
29	Abel et al, 2005	20 long-term-care facilities (United States)	QIO (not specified)	Not reported	Care planning tool Patient/family education tool Communication form to physicians	Training for nursing home staff
30	Baier et al, 2003	29 long-term-care facilities (United States)	QI teams at each site (not specified)	Use of QI approach based on Plan-Study-Act, tailored to each facility	Not reported	Not reported
31	McKeeney, 2008	9 long-term-care facilities (United Kingdom)	Tissue viability service (not specified) 1 wound nurse	Audit tool completed to assess improvement areas Action plan based on outcomes of audit	Not reported	Two formal small group training sessions (at launch) Ongoing support, advice, and education of nursing staff

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
Not reported	Not reported	External facilitation: Periodic visits by QIO Reference cards Wound assessment guide Visual reminder of mobility needs	Not reported	Incidence rate (no significant decrease): *Baseline*: 13.6% *Post*: 10%	Pre/Post data abstracted from medical records to Measure 12 quality indicators *Mixed results* Improvement in 8 out of 12 measures Relationship between improved QI scores and incidence rates
Audit and feedback	Not reported	External facilitation: Individual mentoring for each QI team Collaboration among nursing homes Guidebook for PrU prevention	Not reported	Not reported	12 process measures (data abstracted from records) Did not measure prevalence/incidence rates Measured at baseline and at follow-up after 1 y *Mixed results*: 9 of 12 process measures showed significant improvement (aggregate across all facilities)
Rounding of wound nurse	Tissue viability link-nurse in each long-term-care facility	Manual of pressure relieving equipment	Not reported	Not reported	Audit tool: 9 care processes at baseline and after 8 wk Overall benchmark improved at all sites Action plan to work on problem areas identified through audit

Table 1.
STUDIES DESCRIBING MULTIFACETED PRESSURE ULCER PREVENTION PROGRAMS (LISTED IN ALPHABETICAL ORDER), CONTINUED

Reference	Authors	Setting and Scope	Team	Preparations	Best Practices	Staff Education
						through wound care sister
32	Rosen et al, 2006	Long-term-care facility (United States)	Research team Interacting with administration (not specified)	Not reported	Not reported	4-wk training period prior to intervention (mandatory staff training)

Abbreviations: ARNP, advanced registered nurse practitioner; CMO, Chief Medical Officer; CNO, Chief Nursing Officer; OR, operating room; CC, critical care; ET, enterostomal therapist; CWOCN, certified wound, ostomy and continence nurse; HAPU, hospital-acquired pressure ulcers; QI, quality improvement; QIO, quality improvement organization; TED, thrombo embolic deterrent (stockings); SCD, spinal cord disease; SOS, Save Our Skin (name of campaign).

program.[10,13–15,21,26,27] The roles and responsibilities of skin care champions varied slightly from site to site. Generally, the staff members received additional training in PrU treatment and prevention, and their responsibilities included a combination of the following elements: to introduce the new policies and interventions on the unit, to serve as skin care resource and mentor to coworkers, to serve as liaison between the unit and other parties involved in the improvement efforts, and to participate in the data collection and ongoing process monitoring.

Other Elements. To increase awareness and provide cues to action for consistent and correct implementation of the new clinical practices, several programs utilized audiovisual support and other isolated activities. [9,10,12,13,15,17,18,20–22,24,27,29–32] Examples of these support elements included use of turn clocks, stickers in the patient charts or outside patient rooms to identify patients with PrUs or at risk for developing one, PrU pocket guides and reference cards, theme songs played every 2 hours, penlights for skin assessments, weekly skin care newsletters, posters on the units, and manuals or guidebooks on skin care products, support equipment, or PrU prevention and treatment protocols. Another strategy to ensure continued awareness of the program was the development of acronyms and themes related to the program. The following themes were identified in the reviewed literature: ''Check, Rock & Roll Around the Clock''[9]; PUPPI (Pressure Ulcer Prevention Protocol Interventions)[11]; ''Save Our Skin''[13,27]; SKIN (*S*urfaces, *K*eep the patients turning, *I*ncontinence management, *N*utrition)[18]; and TOE (*T*urn, *O*verlay, *E*levate).[22]

Strategies to Sustain Efforts. It is not evident in the literature if initiatives were discontinued after completion of the project period, or whether all or certain intervention components were continued even after the formal study phase. A number of studies, however, suggested ongoing measurement and reporting of PrU rates as a strategy for ensuring continued awareness of PrU prevention.[9,11,13,14,18,23,24,26] Four studies reported identification of new practice issues as a way for sustaining the momentum of the prevention efforts.[16,18,24,28] Finally, 1 study noted that additional visual and auditory cues were introduced after completion of the study to ensure consistent adherence to PrU prevention protocols.[22]

Clinical Monitoring/Feedback	Skin Care Champions	Other Elements	Sustainability	Pressure Ulcer (PrU) Rates	Care Process Measures
					Evaluation form for training sessions
Weekly feedbacks on staff compliance with training Weekly reports by management to staff of PrU incidence	Not reported	Penlights for skin assessment Caregivers required to wear TAP card $75 if the incidence of PrUs was reduced below the goal $10 for completion of training program	Not reported	Monitored the incidence of PrUs during the 48-wk program *Baseline* HAPU incidence: 28.3% *Intervention* HAPU incidence: 9.3% Reduction of PrU incidence was highest during the 12-wk intervention phase; results were not sustained without the support of the research team	Not reported

Outcomes

PrU Rates. The majority of studies reported positive outcomes from their PrU prevention initiatives; however, *P* values assessing statistical significance were rarely reported. Almost all of the reviewed studies measured PrU prevalence rates before and after implementation of their quality improvement projects.[9–28] Seven studies did not sufficiently describe the results of their prevalence surveys to draw meaningful conclusions.[12,13,18,22,26–28] Eleven studies saw a decrease in prevalence rates over the course of the study period,[9,10,11,14–17,20,21,23,24] whereas 2 programs reported no significant changes.[19,25]

Ten studies reported PrU incidence rates.[9,12,13,18,24–29,32] Eight of these studies reported a decrease in rates between baseline and follow-up; 1 study reported that incidence rates increased between project year 1 and year 4 without statistical significance,[25] and 1 study noted that results could not be sustained during the postimplementation phase.[32]

Care Processes. Process measures were reported by 2 acute care-setting studies[14,16,] and 3 long-term-care-facility studies.[29–31] One acute care-setting study measured the use of a new mattress and implementation of a repositioning schedule and found no significant change in preventive behavior when the use of new support mattresses was not taken into account.[14] The other acute care study reported minimal improvements in some of the measured care processes.[16] Among the 3 multisite long-term-care facility programs, one reported an overall improvement of clinical practice benchmarks across all participating facilities after implementation of the 8-week program.[31] The other 2 multisite studies saw significant improvement in 8 of 12 and 9 of 12 care processes when aggregated across the participating sites.[29,30]

Other Outcomes. Positive outcomes, such as increased staff awareness and knowledge, as well as change in attitudes toward

PRESSURE ULCER TOOLKIT

This literature review was performed as part of the process for developing a toolkit to reduce PrUs using a multifaceted quality improvement approach. The free toolkit is available on the Agency for Healthcare Research and Quality's website: http://www.ahrq.gov/research/ltc/pressureulcertoolkit/.

PrU prevention, were noted in several of the articles. However, reports of these outcomes were mostly anecdotal and were not validated by any formal evaluation.

DISCUSSION

This review showed that there is an array of studies describing the use of multipronged initiatives to prevent PrU development among patients in hospitals and long-term-care facilities. Moreover, many programs reported impressive improvements in PrU prevalence or incidence rates. These results suggest that multifaceted, multidisciplinary programs are effective in preventing PrUs.

A number of approaches were widely used as components of the multipronged approach and are likely to contribute to its success. In preparation of implementation, literature reviews or assessment of the current state of PrU practice was often used to provide a baseline for the data collection and identify areas in need for improvement. Intervention components included the use of a ''bundle'' of best practices for PrU prevention, the reliance on a unit skin care champions, and an emphasis on staff education. Strategies to generate staff enthusiasm and increase awareness and adherence to the best practices, such as turning clocks or skin care newsletters, often were used. The involvement of frontline staff members in all stages of program design and implementation was considered to be essential by many studies to ensure staff engagement, ownership, and dedication. Providing frequent real-time data feedback and giving staff credit for improvement, celebrating success, and stimulating a healthy competition among the units were frequently described as ways of engaging the staff and providing them a sense of pride in their accomplishments. Finally, regular monitoring of charts, weekly or monthly rounding, and root-cause analysis to examine what went wrong when a PrU developed were also used successfully.

Few studies commented on long-term sustainability of the intervention, and there was little in the literature to suggest how improvements could be maintained. Continuous monitoring of PrU rates, the presence of a wound care team or unit champions, and continued formal and informal education seemed to be some of the elements that could positively influence the maintenance of positive outcomes.

Despite the number of studies showing benefit, results must be interpreted with caution. Foremost, the level of evidence is weak. Studies mostly consisted of a longitudinal 1-group pretest-posttest design. They have neither randomization to interventions nor control groups. Description of methods for data collection and analysis was often neglected in the publications. Only 5 studies reported process measures. This makes it difficult to determine whether the interventions contributed to increased staff compliance with new PrU prevention practices. In addition, some of the studies that measured care processes showed that albeit improved, adherence to certain best practices still remained low. One of the studies, for example, found that the proportion of residents with appropriate risk assessment completed within 2 days of admission increased from 2.2% to only 15.3%, whereas the proportion of residents with PrUs that receive weekly skin assessments increased from 12.6% to 32.8%.[29]

Furthermore, the components of the multifaceted programs were not evaluated individually, and it is therefore not possible to determine the impact of each single component. There is also a high likelihood of publication bias. Nearly all published studies were positive in showing a benefit. Given the multitude of interventions, it is not plausible that all programs would work. More likely, those programs that showed a benefit were more likely to be written up and published. Finally, studies generally did not describe or analyze the processes by which the new programs were implemented, the challenges they faced, and how they did overcome them. However, organizational change requires attention not only to the content of the program, but also to the strategies needed to implement the program.

CONCLUSION

Improving PrU prevention remains an important issue for hospital patients. This literature review has identified many components that have consistently been included in successful multifaceted PrU interventions. A review of the studies supports previously reported exemplars of success in PrU reduction initiatives. This includes administrative support with active involvement of clinical staff at the patient care level, bundling of care practices and infusing them into routine care practice, creating systemwide change and communication that is individualized to the institution's culture, making visible the documentation of PrU prevention practices, and regular education of all levels of staff.[33] Prevention practices of risk assessment, pressure redistribution and repositioning, and attention to skin care are common bundle care elements. It appears that the more care practices are incorporated into usual care practices, the more staff are apt to perform them and not see them as ''another task to perform.'' For example, including skin inspection while taking vital signs and reporting a patient's PrU risk assessment status as part of the patient's handoff report are ways for staff to consistently perform suggested prevention interventions. The best outcomes are a result of PrU prevention bundle care practices performed consistently.

Pressure ulcer quality improvement teams that are empowered within their institutions appear to have more success. No one composition of the team has been identified as being best;

each institution must decide what mix of the interdisciplinary team needs to be included in its oversight of PrU reduction initiatives. Pressure ulcer reduction initiatives should be customized and prioritized for the needs of professionals in that institution.[34]

Making too many changes at one time may impact the sustainability of PrU prevention practices. Dahlstrom et al[35] found that also changing to an electronic medical record at the same time as their PrU prevention program decreased gains previously seen. Virani et al[36] have summarized factors that contribute to poor sustainability of evidence-based practices, including inadequate time for teaching new practices to staff, inadequate attention to barriers of acceptance of new practices, and organizational factors such as inadequate resources for equipment/supplies and infrastructure support. They recommend that "sustainability of practice changes therefore requires systematic, thoughtful planning and action to ensure that the changes are embedded into the various knowledge reservoirs in the organization."[36]

The authors agree with Virani et al,[36] who believe that a regular review of research literature and practice guidelines should be used to evaluate an institution's PrU practices. What did not work in one institution might work in another. Facilities that have implemented PrU programs, successfully or not, should be encouraged to rigorously evaluate their programs and publish their results to strengthen the level of evidence.

PRACTICE PEARLS

- The review of the literature supports multi-disciplinary, bundled approaches to PrU prevention.
- No one approach has been identified as being best; institutions should carefully plan and customize PrU prevention programs according to their needs and abilities.
- PrU prevention programs require attention to content, as well as implementation and sustainability strategies.
- Institutions having implemented PrU prevention programs are encouraged to share their strategies and results so that others can learn from the experience.

REFERENCES

1. VanGilder C, Amlung S, Harrison P, Meyer S. Results of the 2008-2009 International Pressure Ulcer Prevalence Survey and a 3-year, acute care, unit-specific analysis. Ostomy Wound Manage 2009;55(11):39-45.
2. HealthGrades. The Sixth Annual HealthGrades Patient Safety in American Hospitals Study. 2009. http://www.healthgrades.com/business/img/PatientSafetyInAmericanHospitalsStudy2009.pdf. Last accessed 02/02/2012.
3. Inpatient Prospective Payment System FY 2009 Final Rule (HAC Section begins on page 39). 2009. http://edocket.access.gpo.gov/2008/pdf/E8-17914.pdf. Last accessed January 25, 2012.
4. Pancorbo-Hidalgo PL, Garcia-Fernandez FP, Lopez-Medina IM, Alvarez-Nieto C. Risk assessment scales for pressure ulcer prevention: a systematic review. J Adv Nurs 2006;54(1):94-110.
5. McInnes E, Bell-Syer SE, Dumville JC, Legood R, Cullum NA. Support surfaces for pressure ulcer prevention. Cochrane Database Syst Rev 2008;(4):CD001735.
6. Rich SE, Shardell, M, Margolis D, Baumgarten M. Pressure ulcer prevention device use among elderly patients early in the hospital stay. Nurs Res 2009;58:95-104.
7. Lyder CH, Preston J, Grady JN, et al. Quality of care for hospitalized Medicare patients at risk for pressure ulcers. Arch Intern Med 2001;161:1549-54.
8. Saliba D, Rubenstein L, Simon B, et al. Adherence to pressure ulcer prevention guidelines: implications for nursing home quality. J Am Geriatr Soc 2003;51(1):56-62.
9. Baldelli P, Paciella M. Creation and implementation of a pressure ulcer prevention bundle improves patient outcomes. Am J Med Qual 2008;23:136-42.
10. Bales I, Padwojski A. Reaching for the moon: achieving zero pressure ulcer prevalence. J Wound Care. 2009;18:137-44.
11. Catania K, Huang C, James P, Madison M, Moran M, Ohr M. Wound wise: PUPPI: the Pressure Ulcer Prevention Protocol Interventions. Am J Nurs 2007;107(4):44-52.
12. Chicano SG, Drolshagen C. Reducing hospital-acquired pressure ulcers. J Wound Ostomy Continence Nurs 2009;36(1):45-50.
13. Courtney BA, Ruppman JB, Cooper HM. Save our skin: initiative cuts pressure ulcer incidence in half. Nurs Manage 2006;37(4):36, 38, 40 passim.
14. De Laat EH, Schoonhoven L, Pickkers P, Verbeek AL, Van Achterberg T. Implementation of a new policy results in a decrease of pressure ulcer frequency. Int J Qual Health Care 2006;18:107-12.
15. Dibsie LG. Implementing evidence-based practice to prevent skin breakdown. Crit Care Nurs Q 2008;31:140-9.
16. Elliott J. Strategies to improve the prevention of pressure ulcers. Nurs Older People 2010;22(9):31-6.
17. Elliott R, McKinley S, Fox V. Quality improvement program to reduce the prevalence of pressure ulcers in an intensive care unit. Am J Crit Care 2008;17:328-34.
18. Gibbons W, Shanks HT, Kleinhelter P, Jones P. Eliminating facility-acquired pressure ulcers at Ascension Health. Jt Comm J Qual Patient Saf 2006;32:488-96.
19. Gunningberg L, Stotts NA. Tracking quality over time: what do pressure ulcer data show? Int J Qual Health Care 2008;20:246-53.
20. Hiser B, Rochette J, Philbin S, Lowerhouse N, Terburgh C, Pietsch C. Implementing a pressure ulcer prevention program and enhancing the role of the CWOCN: impact on outcomes. Ostomy Wound Manage 2006;52(2):48-59.
21. Hopkins B, Hanlon M, Yauk S, Sykes S, Rose T, Cleary A. Reducing nosocomial pressure ulcers in an acute care facility. J Nurs Care Qual 2000;14(3):28-36.
22. LeMaster KM. Reducing incidence and prevalence of hospital-acquired pressure ulcers at Genesis Medical Center. Jt Comm J Qual Patient Saf 2007;33:611-6, 585.
23. McInerney JA. Reducing hospital-acquired pressure ulcer prevalence through a focused prevention program. Adv Skin Wound Care 2008;21:75-8.
24. Sacharok C, Drew J. Use of a total quality management model to reduce pressure ulcer prevalence in the acute care setting. J Wound Ostomy Continence Nurs 1998;25:88-92.
25. Stausberg J, Lehmann N, Kröger K, Maier I, Schneider H, Niebel W. Interdisciplinary decubitus project. Increasing pressure ulcer rates and changes in delivery of care: a retrospective analysis at a university clinic. J Clin Nurs 2010;19(11-12):1504-9.
26. Stoelting J, McKenna L, Taggart E, Mottar R, Jeffers BR, Wendler MC. Prevention of nosocomial pressure ulcers: a process improvement project. J Wound Ostomy Continence Nurs 2007;34:382-8.
27. Young J, Ernsting M, Kehoe A, Holmes K. Results of a clinician-led evidence-based task force initiative relating to pressure ulcer risk assessment and prevention. J Wound Ostomy Continence Nurs 2010;37:495-503.
28. Young ZF, Evans A, Davis J. Nosocomial pressure ulcer prevention: a successful project. J Nurs Adm 2003;33:380-3.
29. Abel RL, Warren K, Bean G, et al. Quality improvement in nursing homes in Texas: results from a pressure ulcer prevention project. J Am Med Dir Assoc 2005;6:181-8.
30. Baier RR, Gifford DR, Lyder CH, et al. Quality improvement for pressure ulcer care in the nursing home setting: the Northeast Pressure Ulcer Project. J Am Med Dir Assoc 2003;4:291-301.

31. McKeeney L. Improving pressure ulcer prevention in nursing care homes. Br J Community Nurs 2008;13(9):S15-6, S18, S20.
32. Rosen J, Mittal V, Degenholtz H, et al. Ability, incentives, and management feedback: organizational change to reduce pressure ulcers in a nursing home. J Am Med Dir Assoc 2006;7:141-6.
33. Lyder CH, Ayello EA. Annual checkup: the CMS pressure ulcer present-on-admission indicator. Adv Skin Wound Care 2009;22:476-84.
34. Ayello EA. Changing systems, changing cultures: reducing pressure ulcers in hospitals. Jt Comm J Qual Patient Saf 2011;37(3):120-2.
35. Dahlstrom M, Best T, Baker C, et al. Improving identification and documentation of pressure ulcers at an urban academic hospital. Jt Comm J Qual Patient Saf 2011;37(3):123-30.
36. Virani T, Lemieux-Charles L, Davis DA, Berta W. Sustaining change: once evidence-based practices are transferred, what then? Healthc Q 2009;12(1):89-96, 2.

For more than 71 additional continuing education articles related to Skin and Wound Care topics, go to NursingCenter.com/CE.

CE Connection

CONTINUING MEDICAL EDUCATION INFORMATION FOR PHYSICIANS

Lippincott Continuing Medical Education Institute, Inc. is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

Lippincott Continuing Medical Education Institute, Inc. designates this journal-based CME activity for a maximum of 1 *AMA PRA Category 1 Credit™*. Physicians should only claim credit commensurate with the extent of their participation in the activity.

PROVIDER ACCREDITATION INFORMATION FOR NURSES

Lippincott Williams & Wilkins, publisher of the *Advances in Skin & Wound Care* journal, will award 2.9 contact hours for this continuing nursing education activity.

LWW is accredited as a provider of continuing nursing education by the American Nurses Credentialing Center's Commission on Accreditation.

This activity is also provider approved by the California Board of Registered Nursing, Provider Number CEP 11749 for 2.9 contact hours. Lippincott Williams & Wilkins is also an approved provider of continuing nursing education by the District of Columbia and Florida #FBN2454.

Your certificate is valid in all states.

The ANCC's accreditation status of Lippincott Williams & Wilkins Department of Continuing Education refers only to its continuing nursing education activities and does not imply Commission on Accreditation approval or endorsement of any commercial product.

CONTINUING EDUCATION INSTRUCTIONS

- Read the article beginning on page 167.
- Take the test, recording your answers in the test answers section (Section B) of the CE enrollment form. Each question has only one correct answer.
- Complete registration information (Section A) and course evaluation (Section C).
- Mail completed test with registration fee to: Lippincott Williams & Wilkins, CE Group, 2710 Yorktowne Blvd, Brick, NJ 08723.
- Within 3 to 4 weeks after your CE enrollment form is received, you will be notified of your test results.
- If you pass, you will receive a certificate of earned contact hours and an answer key. Nurses who fail have the option of taking the test again at no additional cost. Only the first entry sent by physicians will be accepted for credit.
- A passing score for this test is 12 correct answers.
- Nurses: Need CE STAT? Visit **http://www.nursingcenter.com** for immediate results, other CE activities, and your personalized CE planner tool. No Internet access? Call 1-800-787-8985 for other rush service options.
- Questions? Contact Lippincott Williams & Wilkins: 1-800-787-8985.

Registration Deadline: April 30, 2014 (nurses); April 30, 2013 (physicians)

PAYMENT AND DISCOUNTS

- The registration fee for this test is $24.95 for nurses; $22 for physicians.
- Nurses: If you take two or more tests in any nursing journal published by LWW and send in your CE enrollment forms together by mail, you may deduct $0.95 from the price of each test. We offer special discounts for as few as six tests and institutional bulk discounts for multiple tests. Call 1-800-787-8985 for more information.

ORIGINAL RESEARCH ARTICLE

Sexually transmitted infections among HIV-infected heavy drinkers in St Petersburg, Russia

C A Pace MD*, **D Lioznov** MD PhD†, **D M Cheng** ScD*‡, **S E Wakeman** MD§, **A Raj** PhD**, **A Y Walley** MD MSc*, **S M Coleman** MS MPH††, **C Bridden** MA MPH*, **E Krupitsky** MD PhD‡‡ and **J H Samet** MD MA MPH*§§

*Department of Medicine, Section of General Internal Medicine, Clinical Addiction Research and Education (CARE) Unit, Boston University School of Medicine/Boston Medical Center, Boston, MA, USA; †Department of Infectious Diseases and Epidemiology, Pavlov State Medical University, St Petersburg, Russia; ‡Department of Biostatistics, Boston University School of Public Health; §Department of Medicine, Massachusetts General Hospital, Boston, MA; **Division of Global Public Health, Department of Medicine, University of California at San Diego, San Diego, CA; ††Data Coordinating Center, Boston University School of Public Health, Boston, MA, USA; ‡‡St Petersburg Scientific-Research Center of Addictions and Psychopharmacology, Pavlov State Medical University, St Petersburg, Russia; §§Department of Community Health Sciences, Boston University School of Public Health, Boston, MA, USA

Summary: The objective of this study was to estimate the prevalence and identify correlates of four sexually transmitted infections (STIs) among HIV-infected Russians reporting heavy alcohol use and recent unprotected sex, we conducted a cross-sectional analysis of baseline data from the HERMITAGE study. The primary outcome was any current STI, based on urine tests for *Neisseria gonorrhoeae*, *Chlamydia trachomatis* and *Trichomonas vaginalis* and serological testing for infection with *Treponema pallidum*. Data on potential demographic and behavioural predictors of STI were obtained from surveys administered at study entry. Of 682 participants, 12.8% (95% confidence interval [CI] 10.3, 15.3) tested positive for at least one STI. In a multivariable model adjusted for gender, age and marital status, only sex trade involvement over the last three months was significantly associated with an increased odds of STI (adjusted odds ratio [AOR] 2.00, 95% CI 1.13, 3.55). Given that STIs were common in this HIV-infected cohort, and that few patient characteristics predicted STI, the current practice of screening HIV-infected Russians for syphilis alone merits re-evaluation.

Keywords: sexually transmitted infection, HIV, AIDS, STIs, risk behaviour, alcohol drinking, Russia

INTRODUCTION

The HIV epidemic in Russia has expanded rapidly over the last decade and continues to grow, with an estimated one million HIV-infected people in 2010.[1] While the epidemic has been largely driven by injection drug use, the proportion of new infections attributed to sexual risk is increasing.[2,3] This shift has drawn attention to the possible role of sexually transmitted infections (STIs) in facilitating the spread of HIV in the country. Russia experienced rising STI rates in the period preceding its HIV epidemic[4] and STIs are known to increase both susceptibility to HIV and the infectiousness of those with HIV.[5,6]

Studies outside Russia have demonstrated that STIs, particularly gonorrhoea, chlamydia and syphilis, are prevalent in a number of HIV-infected subpopulations.[7–12] However, there are few studies addressing STIs among HIV-infected Russians.[12] Existing data derive from studies based on self-report of STI symptoms, or only a small number of HIV-infected participants.[2,13] These limited data suggest that STIs may make a significant contribution to heterosexual transmission of HIV in Russia: for example, among 32 HIV-infected injecting drug users (IDUs) in one study, 18 (56%) tested positive for infection with either *Chlamydia trachomatis* (herein referred to as chlamydia), *Neisseria gonorrhoeae* (gonorrhoea), *Treponema pallidum* (syphilis), *Trichomonas vaginalis* (trichomoniasis) or herpes simplex virus type-2.[13] Data from a larger sample on the prevalence and correlates of STIs among those infected with HIV in Russia could inform clinical practice as to the appropriate role for STI screening for this group, which, according to national guidelines, is screened routinely only for syphilis, but not for other STIs.

The HERMITAGE (HIV Evolution in Russia – Mitigating Infection Transmission and Alcoholism in a Growing Epidemic) study provided an opportunity to investigate the prevalence and correlates of four STIs among 682 HIV-infected Russians. HERMITAGE is a randomized controlled trial of a behavioural intervention to reduce high-risk sexual activity and substance use among HIV-infected heavy alcohol users who reported recent unprotected sex in St Petersburg. The current study was a cross-sectional, secondary analysis which investigated the prevalence of gonorrhoea, chlamydia, syphilis and trichomoniasis at study entry, and examined which behaviours and demographic characteristics were associated with testing positive for one or more of these STIs.

Correspondence to: Christine A Pace, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor/Boston, MA 02118, USA
Email: Christine.pace@bmc.org

METHODS

Study design and participants

From October 2007 to April 2010, the HERMITAGE study recruited HIV-infected heavy drinkers who reported recent unprotected sex from five inpatient and outpatient HIV and substance use care sites in St Petersburg, Russia. In the clinical settings, research associates approached patients, assessed eligibility, offered participation and conducted assessments. At the non-clinical recruitment site, a needle exchange, research associates gave potential participants information on the study and referred them to one of the clinical sites for eligibility assessment.

Patients were eligible for inclusion in the study if they were 18 years of age or older; HIV-infected; reported anal or vaginal sex without a condom in the past six months; and reported past six month National Institute on Alcohol Abuse and Alcoholism (NIAAA) 'at risk' drinking levels (defined as >14 drinks per week or >4 drinks on a single occasion for men, and >7 per week or >3 on a single occasion, for women), described as 'heavy drinking' herein.[14] Exclusion criteria included cognitive impairment, acute illness precluding participation, pending legal issues which could lead to incarceration or ongoing efforts to conceive.

Procedure

After eligibility assessment, all participants provided written informed consent. Baseline data were collected via: (a) a face-to-face interview with a research associate and (b) a self-administered questionnaire for particularly sensitive questions (e.g. about sexual victimization). A medical chart review was performed for participants recruited from medical settings. All participants were asked to provide a urine and blood sample for STI screening at the baseline assessment. Those testing positive for an STI at baseline were offered treatment.

Interviews were conducted in Russian. Participants were compensated 200 rubles, the equivalent of approximately seven US dollars, for the baseline assessment and received 30 condoms. The HERMITAGE study was approved by the Institutional Review Boards of Boston Medical Center and St Petersburg Pavlov State Medical University.

Measures

Outcome measures

The primary outcome of the current study was a diagnosis of one or more STIs (gonorrhoea, chlamydia, trichomoniasis and/or syphilis) based on biological testing at study entry. Urine specimens were tested for *N. gonorrhoeae*, *C. trachomatis* and *T. vaginalis* using polymerase chain reaction (PCR)-based nucleic acid amplification (Amplisens PCR Kits, Ecoli Ltd, Moscow, Russia) in compliance with quality guidelines by the Federal State Institution of Science, Central Research Institute of Epidemiology, Moscow, Russia).[15,16] Serum was tested for syphilis using the venereal diseases research laboratory test (VDRL, Institute of Vaccines and Sera, St Petersburg, Russia) and an enzyme-linked immunoassay (ELISA) that employs recombinant antigen to detect both IgM and IgG antibodies to *T. pallidum*. (RecombiBest AntiPallidum, ZAO Vector-Best, Novosibirsk, Russia). Subjects were defined as having syphilis when both the ELISA and VDRL were positive and they had not been previously treated for syphilis. Participants with a positive VDRL and negative ELISA were considered false-positives. None of the participants with a positive ELISA and no prior treatment had a negative VDRL.

Potential predictors

Potential predictors of STI were identified from the baseline assessment. Sociodemographic characteristics of interest included age, gender and marital status. Potential health-related predictors included self-report of current antiretroviral therapy or past history of STI, as well as the most recent CD4 count according to a medical record review. Sexual risk behaviours over the last three months were assessed, and included the number of sexual partners; the number of sexual encounters when a condom was not used; buying or selling sex for drugs or money; and, for men, any sexual activity with other men (MSM).[17] In addition, the assessment asked about the use of alcohol or drugs before or during sex in the last 30 days, a lifetime history of sexual victimization and the propensity to seek out novel or risky sexual stimulation, using Kalichman's 11-item sexual sensation-seeking scale. This scale generates scores ranging from 1 to 4, with higher scores reflecting greater sensation seeking.[18] To assess potential alcohol-related predictors of STI, participants were asked about the quantity and frequency of alcohol consumption using a 30-day time-line follow-back.[19] Heavy alcohol use was defined according to the NIAAA definition, described above.[14] Alcohol dependence over the last year was evaluated using the Composite International Diagnostic Interview Short-Form (CIDI-SF).[20] Drug-related variables included any use of heroin, prescription analgesics, marijuana, sedatives, tranquilizers or stimulants (including amphetamines and cocaine) over the last year, drug dependence according to the CIDI-SF and past history of injecting drug use.

Data analysis

Descriptive statistics were used to characterize the study sample at baseline, both overall and by STI status. The frequency and proportion of STIs (i.e. testing positive by biological testing for any of 4 specific STIs), and of any history of STI by self-report were also described. To identify factors associated with testing positive for an STI, a series of logistic regression models were built using the following manual, iterative model building approach. First, unadjusted logistic regression models for each factor of interest were fit. Factors with a P value of less than 0.15 were then included together in a single multivariable model. In addition, covariates considered to be important potential confounders (i.e. gender, age and marital status) were forced into the multivariable model. Factors in this multivariable model with a P value greater than 0.15 were removed one at a time. Finally, factors not selected based on the initial unadjusted analyses were included one at a time in the current multivariable model to assess their importance in the presence of other variables. The continuous variables of age, sexual sensation seeking score and number of unprotected sexual encounters were included as tertiles. CD4 count was included as a dichotomous variable ($\geq$350 versus $<$350 cells/mm^3) as was the number of sexual partners ($<$2 versus $\geq$2), which is consistent both with the distribution of our data and the standard approach for measuring multiple partnering over fairly short time periods.[21] To minimize the

potential for collinearity, we assessed the correlation between all pairs of independent variables and verified that no pair of variables included in the same regression model was highly correlated (i.e. $r > 0.40$). Due to a large number of missing values for CD4 cell count (30% of the sample), we used an indicator variable to create a category for missing values for this variable.

Secondary, exploratory analyses were also conducted to assess gender as a possible effect modifier. This analysis aimed to test the hypotheses that younger age may be a more important risk factor for STIs among women than men, and that the effects of certain behaviours on STI risk may also differ by gender.[20] Potential interactions between gender and four factors hypothesized to have differential effects on STI risk among women and men (age, any alcohol use before or during sexual activity, sex trade involvement and past year cocaine use) were evaluated.

All reported P values were two-tailed, and a P value of less than 0.05 was considered statistically significant. Analyses were performed using SAS software (version 9.1; SAS Institute, Cary, NC, USA).

RESULTS

As shown in Figure 1, 921 individuals were screened for eligibility. Of the 190 ineligible individuals, 110 did not meet the alcohol-related criteria and 134 did not meet the sex risk criteria. Seven hundred of the 731 eligible individuals (96%) agreed to participate. Among these, 682 had available STI results, and were included in the current analysis.

Demographic, health-related and behavioural characteristics of the study cohort are shown in Table 1. In general, the sample was young, 40% were women, and, reflecting study eligibility criteria, had a high prevalence of heavy alcohol use and unprotected sex.

The prevalence of any STI among the 682 participants was 12.8% (95% confidence interval [CI] 10.3, 15.3; Table 2). The prevalence appeared similar in women and men. Only 1.9% of all participants ($n = 13$) had more than one STI. Chlamydia was diagnosed among 40 participants (5.8%); it was more prevalent among men than women (7.3% of men versus 3.6% of women). Trichomoniasis was also common with 37 participants (5.4%) testing positive; prevalence was 8.2% among women and 3.4% among men. Fourteen participants (2.0%) had gonorrhoea and 13 (1.9%) met the case definition of untreated syphilis infection. About half of participants (47.7%) reported a past diagnosis of an STI, including syphilis, gonorrhoea, chlamydia, trichomoniasis, genital warts or herpes.

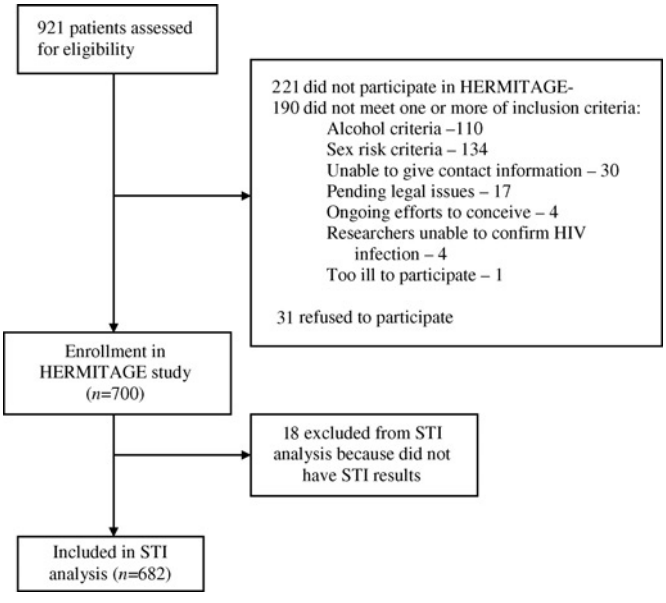


Figure 1 **Enrollment of participants in the HERMITAGE study and inclusion in the current analysis of sexually transmitted infection (STI) data**

Bivariate analyses

Based on the preliminary bivariate analyses, seven independent variables met criteria for entry into the initial multivariable model: the number of unprotected sexual encounters (odds ratio [OR] 0.61, 95% CI 0.34, 1.10 for the middle versus lowest tertile; OR 1.13, 95% CI 0.67, 1.91 for the highest versus lowest tertile); buying or selling sex (OR 1.87, 95% CI 1.08, 3.24); any history of STI (OR 1.49, 95% CI 0.95, 2.34); MSM (OR 2.82, 95% CI 0.86, 9.20); using drugs or alcohol before sex (OR 1.50, 95% CI 0.95, 2.37); using alcohol before sex (OR 1.53, 95% CI 0.98, 2.41); and past year cocaine use (OR 1.92, 95% CI 1.01, 3.62). Using drugs *or* alcohol before sex was not included in the multivariable model due to its potential collinearity with using alcohol before sex, and past IDU was not included due to its potential collinearity with past year heroin use. None of the other potential predictors described in the Methods section met criteria for inclusion in the multivariable model.

Final multivariable model

Using the model building approach outlined in the Methods, a final multivariable model was developed and is shown in Table 3. In this model, buying or selling sex was associated with a two-fold increased odds of having an STI (adjusted odds ratio [AOR] 2.00, 95% CI 1.13, 3.55). No other factors were significantly associated with testing positive for an STI, though there were notable, non-significant increases in the odds of an STI among those reporting a past STI (AOR 1.46, 95% CI 0.91, 2.35) and among MSM compared with men reporting sex with women only (AOR 2.72, 95% CI 0.78, 9.49).

In secondary, exploratory analyses, there was no statistically significant interaction between gender and any of the following variables: age, sex trade, cocaine use in the past year or alcohol use before or during sexual activity.

DISCUSSION

The prevalence of STIs in the HERMITAGE cohort of 682 HIV-infected heavy drinkers in Russia was 13%. This is comparable with what has been described in HIV-infected populations in other countries.[12] Many in this group reported past injecting drug use, which, consistent with the epidemiology of HIV in Russia, likely led to their HIV infection. However, sexual transmission of HIV is a growing concern in Russia,[2] and the substantial prevalence of STIs in this sexually active cohort indicates a likely role for such infections in facilitating HIV transmission to others.

Table 1 **Baseline characteristics of HIV-infected heavy drinkers with recent unprotected sex in St Petersburg**

Characteristic	*n* (%), total (*n* = 682)	*n* (%), any STI (*n* = 87)	*n* (%), no STI (*n* = 595)
Demographics			
Male	408 (60.0)	48 (55.2)	360 (60.5)
Mean age (range, std)	30.0 (18–57, 5.2)	29.6 (21–44, 5.0)	30.1 (18–57, 5.2)
College/university educated	391 (57.3)	49 (56.3)	342 (57.4)
Currently married or cohabitating	244 (35.8)	30 (34.5)	214 (36.0)
Employed	498 (73.0)	57 (65.5)	399 (67.0)
Ever incarcerated	260 (38.1)	26 (29.9)	234 (39.3)
Health characteristics			
CD4≥350 cells/mm^3 (*n* = 479)*	249 (52.0)	28 (43.1)	221 (53.4)
Currently on ARV therapy	117 (17.2)	11 (12.6)	106 (17.8)
Sexual risk behaviours†			
Two or more partners	181 (26.5)	25 (28.7)	156 (30.1)
Inconsistent condom use‡	518 (76.2)	66 (75.9)	452 (76.2))
Number unprotected sexual encounters (median, IQR)	5 (1–18)	5 (1–26)	5 (1–17)
MSM§	14 (3.4)	4 (4.6)	10 (1.7)
Bought or sold sex for drugs or money	102 (15.0)	20 (23.0)	82 (13.8)
Drugs/alcohol before sex, past 30 days	339 (49.9)	51 (58.6)	288 (48.6)
Substances used before/during sex, past 30 days			
Alcohol	281 (41.4)	44 (50.6)	237 (40.0)
Heroin	149 (21.9)	17 (19.5)	132 (22.3)
Stimulants	40 (5.9)	8 (9.2)	32 (5.4)
Cannabis	35 (5.2)	3 (3.5)	32 (5.4)
Sexual sensation-seeking score (tertiles)			
1–1.750	236 (34.7)	31 (36.1)	205 (34.5)
1.875–2.375	223 (32.8)	27 (31.4)	196 (32.9)
2.50–4.0	222 (32.6)	28 (32.6)	194 (32.6)
Any history of sexual victimization	231 (33.9)	31 (35.6)	200 (33.6)
Alcohol and drug use			
Heavy drinking, last 30 days	554 (81.2)	72 (82.8)	482 (81.0)
Ever injected drugs	566 (83.0)	70 (80.5)	496 (83.4)
Alcohol dependence, last year	434 (63.6)	57 (65.5)	377 (63.4)
Drug use in last year			
Sedative or tranquilizer	256 (37.6)	38 (43.7)	218 (36.7)
Amphetamines or stimulants	213 (31.3)	27 (31.0)	186 (31.3)
Analgesics or prescribed painkillers	227 (33.3)	30 (34.5)	197 (33.1)
Inhalants	15 (2.2)	3 (3.5)	12 (2.0)
Marijuana	319 (46.8)	38 (43.7)	281 (47.2)
Cocaine, crack or free base	68 (10.0)	14 (16.1)	54 (9.1)
LSD/other hallucinogens	53 (7.8)	7 (8.14)	46 (7.7)
Heroin	432 (63.4)	51 (58.6)	381 (64.1)
Drug dependence, last year	425 (62.3)	50 (57.5)	375 (63.0)

std = standard deviation; STI = sexually transmitted infection; IQR = interquartile range; ARV = antiretroviral; MSM = men who have sex with men; LSD = lysergic acid diethylamide 'acid'

*CD4 count was available for 479 participants, of whom 28 had STI and 221 did not have STI. Percentages shown use 479, 65 and 414 as the denominators

†All items refer to behaviours over the last three months, except where noted

‡At least one episode of sex without a condom in the last three months

§Percentages in this row reflect the proportion among men only

Table 2 **Prevalence of sexually transmitted infections (STIs)**

	N (%) among total (*n* = 682) 95% CI
Tested positive for any STI*	87 (12.8) (10.3–15.3)
Specific STIs: (*n* = 682)	
Gonorrhoea	14 (2.0) (1.0–3.4)
Chlamydia	40 (5.8) (4.0–7.5)
Trichomoniasis	37 (5.4) (3.6–7.0)
Syphilis	13 (1.9) (1.0–3.2)
Tested positive for more than one STI	13 (1.9) (1.0–3.2)
Self report of lifetime STI	334 (47.7) (44.0–51.4)

*Gonorrhoea, chlamydia, trichomoniasis or syphilis

It is thus important to examine factors that may be associated with testing positive for an STI in this subgroup of HIV-infected Russians. In this study, the only variable that was significantly predictive of having an STI was buying or selling sex over the last three months, which doubled the odds of STI. The association between selling sex and STIs has been noted in many other studies, including several from Russia,[22–24] where sex work is an important risk factor for HIV itself.[23] However, there have been fewer data on the association between sex trade involvement and STIs from HIV-infected cohorts, a group whose risk behaviours may be different from those without HIV.[7,11] Also of note in the current study was the finding that despite the relationship between sex trade and STIs, there was no relationship between testing positive for an STI and either the number of sexual partners or the number of unprotected sexual encounters. Such a pattern has been observed in at least one other Russian study,[24] suggesting that other, unmeasured sex risk factors (for example, being part of a particular sexual network with a high prevalence of STIs) may help account for STI risk among those who buy and sell sex in Russia.

Table 3 **Characteristics associated with STI in the final multivariable model**

Characteristic	Adjusted OR (95% confidence interval)	*P* value
Age (years)		
18–27	Referent	0.78
28–31	0.95 (0.54, 1.67)	
32–57	0.84 (0.47, 1.50)	
Gender		
Male	0.88 (0.54, 1.43)*	0.60
Marital status		
Married or living with partner	0.92 (0.54, 1.43)	0.76
Number of unprotected sexual encounters† (tertiles)		
1–2	1.00	0.10
3–14	0.59 (0.32, 1.09)	
15–483	1.12 (0.64, 1.94)	
Bought or sold sex in past 3 months	2.00 (1.13, 3.55)	0.02
Past history of STI	1.46 (0.91, 2.35)	0.12
Men who have sex with men	2.72 (0.78, 9.49)‡	0.12

STI = sexually transmitted infection
*Odds ratio (OR) represents comparison of men who report sex with women only versus women
†Over the past three months
‡OR is comparison of men who report any sex with men versus men who report sex with women only, over the last three months

Despite the small number (14, 3.4%) of men who reported MSM, this variable nearly achieved statistical significance as a predictor of STI and had a large magnitude of association with an OR of 2.72. Particularly given that the number of MSM may be an underestimate due to underreporting of stigmatized behaviour in Russia,[25] the finding may be clinically important.

Several independent variables surprisingly did not appear to have strong associations with STIs, most notably age and gender. In many other studies, including some of HIV-infected cohorts, age under 25 years and female gender have been found to be strongly associated with STI.[7,10,23,24,26–28] The impact of both age and gender on overall STI rates in the current study could have been attenuated by the high-risk behaviour exhibited by both men and women who participated. In addition, given that some studies have shown particularly high rates of STIs among 15–19 year olds,[27] the fact that all participants were 18 and over may have lessened the particular effect of age as a predictor in this cohort. The expected effect of gender would also not be seen if the men in this study were involved in disproportionately high-risk sexual activity, such as MSM, that they underreported. In other words, if MSM (or another high-risk, potentially under-reported behaviour among men) was in reality more common than reflected in the data, residual confounding would prevent us from gauging the true impact of gender alone on STI risk in multivariable analysis.

Patterns of alcohol use and the number of unprotected sexual encounters were not significantly associated with having an STI. The study's selection criteria meant that all participants reported heavy drinking and at least one episode of unprotected sex over the last six months. In this context, it may be understandable that neither increasing severity of the alcohol-related problem nor an increasing number of unprotected sexual encounters over the last three months was associated with increased STI risk.

The study also found that neither users of particular drugs, nor those with drug dependence were more likely to have STIs, in contrast to findings from several past studies in both HIV-infected and uninfected populations in the USA in which drug use, particularly marijuana and stimulant use, were risk factors for STI.[7,11,29] This raised the possibility that drug use had a differential effect on STI risk among men compared with women, as was seen in a recent study from India in which alcohol was more strongly associated with condom non-use among men (perhaps because men had more control than women over the decision to use condoms).[30] A secondary analysis was conducted in the current study, to investigate whether a similar process could have occurred for drug use and STI risk in the HERMITAGE cohort. There was, however, no significant interaction seen between gender and cocaine use or alcohol use before sex.

Future studies should investigate the prevalence of STIs in other groups of HIV-infected Russians. Should the current findings be replicated, the data may have implications for STI screening practices in Russia, where current guidelines (in contrast to those issued by the United States Center for Disease Control and Prevention)[31] recommend that HIV-infected patients be screened for syphilis alone. HIV-infected Russians typically receive this screening from their HIV providers, while public STI clinics are available for treatment of symptomatic STIs. Given that STIs are often asymptomatic, many cases likely go untreated with this approach. Our analysis suggests furthermore that screening for STIs only among HIV-infected patients who report certain behavioural risk factors may also be insufficient, since few factors predicted STI in the current cohort, and those factors that were associated with STI (sex trade and perhaps MSM) are stigmatized behaviours that patients may not consistently report to providers. Future research could thus help clarify whether screening all HIV-infected Russians is indeed the optimal strategy, as our data suggest it to be, and whether screening is needed for all four of these treatable STIs or a subset of them.

Our study has several limitations. Rectal specimens, as opposed to urine samples alone, could have provided a more complete assessment for gonorrhoea, chlamydia and trichomoniasis. The lower prevalence of chlamydia among women compared with men additionally raises the possibility of reduced sensitivity of the urine assay among women. Unfortunately, the available peer-reviewed report on the performance characteristics of the Russian chlamydia assay lists a sensitivity for urine specimens (92.3%) that was determined from male specimens alone; only cervical and vaginal specimens were examined from women (sensitivity 87.2%) and these were not possible to obtain in the context of HERMITAGE.[16] Finally, cohort eligibility criteria limited our ability to discern a relationship between alcohol and STIs, or condom use and STIs, and may also limit the study's generalizability to other HIV-infected Russians. However, these behaviours appear to be common among HIV-infected people in Russia; of the 921 HIV-infected patients who were screened for the study, only 110 did not meet the alcohol criteria and 134 failed to meet the sex risk criteria.

In summary, STIs (gonorrhoea, chlamydia, trichomoniasis and syphilis) were common among HIV-infected Russian drinkers who reported recent unprotected sex. Multivariable analyses identified only sex trade involvement as a significant predictor of STI. Future research should investigate the prevalence of these four common, treatable STIs in a broader sample of HIV-infected Russians, to determine whether routine screening for all of these STIs, rather than syphilis

alone, in this population should be a component of HIV prevention efforts in Russia.

ACKNOWLEDGEMENTS

CAP, DL, SEW, AYW, AR, EK and JHS conceived the research question and designed the analysis. JHS obtained funding for the parent HERMITAGE study. DMC and SMC conducted the statistical analysis. All authors participated in interpretation of data. All authors participated in preparation of the manuscript and have approved the final draft for submission. The authors wish to acknowledge the participants in the HERMITAGE study, as well as the clinical and administrative study staff, particularly Elena Blokhina, Victoria Churchill and the Data Coordinating Center at Boston University. In addition, we are grateful to Aseem Mehta for assistance with the literature review, and Guillermo Madico for advice on STI laboratory testing. Support for this work came from the National Institute of Alcohol Abuse and Alcoholism (NIAAA; R01AA016059 and K24-AA015674), the National Institute of Drug Abuse (NIDA; R25DA13582) and the National Institute of Allergy and Infectious Diseases (NIAID; T32 AI052074-06A2).

REFERENCES

1 UNAIDS. *Fact Sheet: Russian Federation*. 2010. See http://www.unaids.org/en/regionscountries/countries/russianfederation. (last checked 11 January 2011)
2 Burchell AN, Calzavara LM, Orekhovsky V, Ladnaya NN, Russian HIV Response Network. Characterization of an emerging heterosexual HIV epidemic in Russia. *Sex Transm Dis* 2008;**35**:807–13
3 Burruano L, Kruglov Y. HIV/AIDS epidemic in Eastern Europe: recent developments in the Russian Federation and Ukraine among women. *Gend Med* 2009;**6**:277–89
4 Borisenko KK, Tichonova LI, Renton AM. Syphilis and other sexually transmitted infections in the Russian Federation. *Int J STD AIDS* 1999;**10**:665–8
5 Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. *Sex Transm Infect* 1999;**75**:3–17
6 Hayes R, Watson-Jones D, Celum C, van de Wijgert J, Wasserheit J. Treatment of sexually transmitted infections for HIV prevention: end of the road or new beginning? *AIDS* 2010;**24**(Suppl. 4):S15–26
7 Kalichman SC, Rompa D, Cage M. Sexually transmitted infections among HIV seropositive men and women. *Sex Transm Infect* 2000;**76**:350–4
8 McClelland RS, Lavreys L, Katingima C, *et al.* Contribution of HIV-1 infection to acquisition of sexually transmitted disease: a 10-year prospective study. *J Infect Dis* 2005;**191**:333–8
9 Phipps W, Stanley H, Kohn R, Stansell J, Klausner JD. Syphilis, chlamydia, and gonorrhea screening in HIV-infected patients in primary care, San Francisco, California, 2003. *AIDS Patient Care STDS* 2005;**19**:495–8
10 Guthrie BL, Kiarie JN, Morrison S, *et al.* Sexually transmitted infections among HIV-1-discordant couples. *PLoS One* 2009;**4**:e8276
11 Mayer KH, O'Cleirigh C, Skeer M, *et al.* Which HIV-infected men who have sex with men in care are engaging in risky sex and acquiring sexually transmitted infections: findings from a Boston community health centre. *Sex Transm Infect* 2010;**86**:66–70
12 Kalichman SC, Pellowski J, Turner C. Prevalence of sexually transmitted co-infections in people living with HIV/AIDS: systematic review with implications for using HIV treatments for prevention. *Sex Transm Infect* 2011;**87**:183–90
13 Abdala N, Krasnoselskikh TV, Durante AJ, Timofeeva MY, Verevochkin SV, Kozlov AP. Sexually transmitted infections, sexual risk behaviors and the risk of heterosexual spread of HIV among and beyond IDUs in St Petersburg, Russia. *Eur Addict Res* 2008;**14**:19–25
14 USDHHS. *The Physicians' Guide to Helping Patients with Alcohol Problems*. US Department of Health and Human Services, National Institute on Alcohol Abuse and Alcoholism, 1995
15 Shipitsyna E, Zolotoverkhaya E, Hjelmevoll SO, *et al.* Evaluation of six nucleic acid amplification tests used for diagnosis of *Neisseria gonorrhoeae* in Russia compared with an international strictly validated real-time porA pseudogene polymerase chain reaction. *J Eur Acad Dermatol Venereol* 2009;**23**:1246–53
16 Shipitsyna E, Zolotoverkhaya E, Agne-Stadling I, *et al.* First evaluation of six nucleic acid amplification tests widely used in the diagnosis of *Chlamydia trachomatis* in Russia. *J Eur Acad Dermatol Venereol* 2009; **23**:268–76
17 Kalichman SC, Rompa D, Cage M, *et al.* Effectiveness of an intervention to reduce HIV transmission risks in HIV-positive people. *Am J Prev Med* 2001;**21**:84–92
18 Kalichman SC, Johnson JR, Adair V, Rompa D, Multhauf K, Kelly JA. Sexual sensation seeking: scale development and predicting AIDS-risk behavior among homosexually active men. *J Pers Assess* 1994; **62**:385–97
19 Sobell LC, Sobell MB. Timeline followback: a technique for assessing self-reported alcohol consumption. In: Litten RZ, Allen JP, eds. *Measuring Alcohol consumption: Psychosocial and Biological Methods Totowa*. New Jersey: The Humana Press, 1992
20 Robins LN, Wing J, Wittchen HU, *et al.* The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988;**45**:1069–77
21 Kalichman SC, Ntseane D, Nthomang K, Segwabe M, Phorano O, Simbaya LC. Recent multiple sexual partners and HIV transmission risks among people living with HIV/AIDS in Botswana. *Sex Transm Infect* 2007;**83**:371–5
22 Karapetyan AF, Sokolovsky YV, Araviyskaya ER, Zvartau EE, Ostrovsky DV, Hagan H. Syphilis among intravenous drug-using population: epidemiological situation in St Petersburg, Russia. *Int J STD AIDS* 2002;**13**:618–23
23 Shakarishvili A, Dubovskaya LK, Zohrabyan LS, *et al.* Sex work, drug use, HIV infection, and spread of sexually transmitted infections in Moscow, Russian Federation. *Lancet* 2005;**366**:57–60
24 Platt L, Rhodes T, Judd A, *et al.* Effects of sex work on the prevalence of syphilis among injection drug users in 3 Russian cities. *Am J Public Health* 2007;**97**:478–85
25 Amirkhanian YA, Kelly JA, Kirsanova AV, *et al.* HIV risk behaviour patterns, predictors, and sexually transmitted disease prevalence in the social networks of young men who have sex with men in St Petersburg, Russia. *Int J STD AIDS* 2006;**17**:50–6
26 Miller WC, Ford CA, Morris M, *et al.* Prevalence of chlamydial and gonococcal infections among young adults in the United States. *JAMA* 2004;**291**:2229–36
27 Datta SD, Sternberg M, Johnson RE, *et al.* Gonorrhea and chlamydia in the United States among persons 14 to 39 years of age, 1999 to 2002. *Ann Intern Med* 2007;**147**:89–96
28 Kourbatova EV, Akovbyan VA, Chesson HW, *et al.* Assessment of the routine, occupation-based gonorrhea and syphilis screening program in Moscow, Russia: an analysis of sexually transmitted infection prevalence and cost-effectiveness. *Sex Transm Dis* 2008;**35**:453–60
29 Tetrault JM, Fiellin DA, Niccolai LM, Sullivan LE. Substance use in patients with sexually transmitted infections: results from a national U.S. survey. *Am J Addict* 2010;**19**:504–9
30 Samet JH, Pace CA, Cheng DM, *et al.* Alcohol use and sex risk behaviors among HIV-infected female sex workers (FSWs) and HIV-infected male clients of FSWs in India. *AIDS Behav* 2010;**14**(Suppl. 1):S74–83
31 CDC. *Sexually Transmitted Diseases Treatment Guidelines*. 2010. See http://www.cdc.gov.ezproxy.bu.edu/mmwr/preview/mmwrhtml/rr5912a1.htm. (last checked 15 June 2011)

(Accepted 12 March 2012)

Adherence to Conservative Management Recommendations for Abnormal Pap Test Results in Adolescents

Rebecca B. Perkins, MD, MSc, Jennifer R. Jorgensen, MD, MPH, Molly E. McCoy, MPH, Sharon M. Bak, MPH, Tracy A. Battaglia, MD, MPH, and Karen M. Freund, MD, MPH

OBJECTIVE: To investigate whether the 2006 American Society for Colposcopy and Cervical Pathology guidelines for conservative management of minimally abnormal Pap test results (atypical squamous cells of undetermined significance, human papillomavirus–positive, and low-grade squamous intraepithelial lesions) and moderate dysplasia (cervical intraepithelial neoplasia 2) in adolescents 1) resulted in fewer colposcopies and loop electrosurgical excision procedures (LEEPs) in adolescents or 2) resulted in unintended treatment changes in older age groups.

METHODS: We analyzed data from 1,806 women aged 18 years and older attending one of six community health centers who were diagnosed with abnormal Pap test results between January 1, 2004, and December 31, 2008. We used multivariable logistic regression to examine treatment differences in women with minimally abnormal Pap test results before and after guideline changes. Variables included date of abnormality, site of care, race or ethnicity, language, and insurance type. We used Fisher exact tests to examine rates of LEEP in patients with moderate dysplasia before and after guideline publication.

RESULTS: Among 206 women aged 18–20 years, rates of colposcopy after a minimally abnormal Pap test result decreased from 78% (n=102) to 45% (n=34) after guideline changes ($P<.001$). Colposcopy among women over age 21 (n=1,542) remained unchanged (greater than 90%). Multivariable logistic regression indicated that both date of abnormality and site of care were associated with colposcopy referral. After guideline changes, management of moderate dysplasia with LEEP in women aged 18–23 decreased from 55% to 18% ($P=.04$); rates remained stable in women ages 24 and older (70% compared with 74%; $P=.72$).

CONCLUSION: Health care providers quickly adopted new conservative management guidelines for low-income, minority adolescents, which may reduce preterm deliveries in these high-risk populations.

(Obstet Gynecol 2012;119:1157–63)
DOI: 10.1097/AOG.0b013e31824e9f2f

LEVEL OF EVIDENCE: II

Although persistent human papillomavirus (HPV) infection causes nearly all cases of cervical dysplasia and cervical cancer,[1] many mild and moderate dysplasias resolve without treatment in adolescent women.[2–4] Before 2007, excision of moderate dysplasia was recommended for all women.[5] However, excision procedures such as loop electrosurgical excision procedures (LEEPs) have been linked to adverse obstetric outcomes including preterm delivery and preterm premature rupture of membranes.[6,7] These risks may be compounded in low-income and minority women, who have higher rates of both HPV infection and preterm delivery than the general population.[8–10] In light of the evidence showing high rates of regression of cervical dysplasia in adolescents and increased risks of adverse obstetric outcomes after

From the Department of Obstetrics and Gynecology, Boston University School of Medicine, and the Boston University School of Public Health, Boston, Massachusetts.

Supported by an American Cancer Society Mentored Research Scholar Grant (MRSG-09-151-01) and the Boston Patient Navigation Research Program (NIH grant U01CA116892).

The authors thank Drs. Richard Kalish, James Taylor, Stephen Tringale, Patrick Egan, Nisha Thakrar, and Ms. Barbara Lottero for their involvement in study design, recruitment, and implementation at their respective clinical sites; Dr. Rebecca Silliman for her review of manuscript drafts; and Dr. Timothy Heeren for his help with statistical analysis.

Presented at the Annual Clinical Meeting of the American College of Obstetricians and Gynecologists, May 5–9, 2012, San Diego, California.

Corresponding author: Rebecca B. Perkins, MD, MSc, 85 E. Concord Street 6th Floor, Department of Obstetrics and Gynecology, Boston University School of Medicine, Boston, MA 02118; e-mail: Rebecca.perkins@bmc.org.

Financial Disclosure
The authors did not report any potential conflicts of interest.

ISSN: 0029-7844/12

LEEPs, the American Society for Colposcopy and Cervical Pathology published new clinical practice guidelines in 2006 recommending repeated Pap testing instead of colposcopy for females younger than 21 years old with minimally abnormal Pap test results, and repeated colposcopy instead of LEEPs for management of moderate dysplasia.[11]

Previous research indicated poor health care provider adherence to published guidelines in the area of cervical cancer screening.[12,13,15,16] Using a database that captured the follow-up of abnormal Pap test results in women aged 18 and older in six independent community health centers between 2004 and 2008, we investigated whether conservative management recommendations in the 2006 American Society of Colposcopy and Cervical Pathology consensus guidelines[11,14] resulted in less aggressive management of minimally abnormal Pap test results and moderate dysplasia in adolescent women, and whether the guidelines had unintended treatment consequences in older age groups.

MATERIALS AND METHODS

Our study is a secondary data analysis of the Boston Patient Navigation Research Program. The Patient Navigation Research Program was a multi-site national trial that evaluated the effects of patient navigation, defined as providing support and guidance to vulnerable persons with cancer or abnormal cancer screening tests, on timeliness and quality of care received. The Boston Patient Navigation Research Program is a community-based intervention implemented at six of Boston's neighborhood community health centers, which serve a population of approximately 50,000 low-income and minority women ages 18 and older.[17,18] Eligible participants for the cervical cancer screening arm of the Boston Patient Navigation Research Program, which constituted the total patient pool for the current study, included nonpregnant women 18 years of age and older with a cervical cancer screening abnormality at one of the participating community health centers. Patient charts were reviewed up until the point at which participants reached a definitive diagnosis of cancer or noncancer or were lost to follow-up. Patient and clinic data were collected from electronic medical records between January 1, 2004, and December 31, 2008. New recommendations regarding the management of abnormal Pap test results were published in October 2007, midway through the Boston Patient Navigator Research program study period, creating the opportunity for a natural experiment comparing management before new guidelines (January 1, 2004, through October 31, 2007) with management after the publication of new guidelines (November 1, 2007, through January 1, 2009). This study was approved by the Boston University School of Medicine Institutional Review Board.

Patient race and ethnicity data were collected at the time of clinical registration. For those with missing race and ethnicity, race was imputed using language and birth country using a cross-tab table in SAS. Five race-ethnicity mutually exclusive categories were used: white, black, Hispanic, Asian, and other. Patient age was collected from the medical record at the time of abnormal Pap test result and reported in years. Patient language was collected at the time of clinical registration. For those missing language, these data were pulled from the text of the participant's medical records. Insurance status was collected at the time of registration. Three insurance categories were used in analyses: public, private, and uninsured. The index abnormality for each patient was defined as the abnormal cervical cancer screening test that made the patient eligible for the study, and it included minimally abnormal Pap test results (atypical squamous cells of undetermined significance [ASC-US]) positive for HPV infection: ASC-US HPV+ and low-grade squamous intraepithelial lesions (LSIL) as well as high-grade squamous intraepithelial lesions. Data were collected during two distinct time periods: from January 1, 2004, through December 31, 2005, and from January 1, 2007, through December 31, 2008. The first time period represents baseline data collection and the second represents the intervention portion of the patient navigation study. The patient navigation intervention did not seek to address whether or not health care providers adhered to new guidelines.

We examined 1) whether treatment of women aged 18–20 years with minimally abnormal Pap test results (ASC-US HPV+ and LSIL) changed in response to new guidelines, specifically whether they were more likely to receive repeat Pap tests instead of colposcopy after guideline changes (defined as before or after October 31, 2007); 2) whether treatment of women ages 21 and older with minimally abnormal Pap test results (ASC-US HPV+ and LSIL) was inadvertently affected by the guidelines, specifically whether they were also more likely to be treated with repeated Pap test instead of colposcopy; 3) whether treatment of women ages 18–20 years with moderate dysplasia (cervical intraepithelial neoplasia [CIN] grade 2) changed in response to new guidelines, specifically whether they were more likely to be treated with repeated colposcopy instead of LEEPs;

and 4) whether treatment of women ages 21 and older with moderate dysplasia (CIN 2) was inadvertently affected by new guidelines, specifically whether they were also more likely to be treated with repeated colposcopy instead of LEEPs. To examine the outcomes related to the management of minimally abnormal Pap test results (ASC-US HPV+ and LSIL), patients were stratified by age into three groups: age 18–20, age 21–25, and age 26 and older. Within each age strata, we first performed univariable analyses to determine the relationships between sociodemographic and clinic characteristics of patients and their receipt of repeated Pap test or colposcopy. Because of the clinical site differences suggested in univariable analyses, we used multivariable logistic regression to examine adjusted effects of the guidelines, patient characteristics, and clinic differences on study outcome. All variables were considered for inclusion in multivariable models. In addition, generalized estimating equation analysis was performed to estimate the effect of patient-level characteristics on adherence to guideline change while controlling for the clustering effect of clinical site of care.

To assess the outcomes related to the management of moderate dysplasia (CIN 2), we analyzed differences in the completion of LEEPs before and after guideline changes using the Fisher exact and χ^2 tests. Although guidelines recommend repeated colposcopy for women with moderate dysplasia (CIN 2) younger than age 21 and LEEPs for women ages 21 and older, literature on regression of lesions includes women up through age 25,[6,7] and some health care providers may choose to defer LEEPs in nulliparous patients older than age 21. Therefore, we performed separate analyses on women ages 18–20, 21–23, 24–25, and 26 years and older. We found similar changes in the rates of LEEPs before and after new guidelines in the 18–20 and 21–23 age groups but different treatment patterns in women older than age 23. Because of similar management outcomes, women ages 18–23 were compared with women ages 24 and older in the final analysis. Analyses were conducted with SAS 9.1.

RESULTS

A total of 1,806 women were included in the analysis; 264 (15%) were aged 18–20 years, and 1,542 aged 21 and older. Demographic characteristics before and after guideline changes were similar with the exception of a larger proportion of 18- to 20-year-old patients before the guideline changes (15% compared with 13%) and a larger proportion of patients from smaller health centers before the guideline changes (17% compared with 7% from Health Center F) (Table 1).

Table 1. Demographics of Patients Whose Abnormal Pap Test Result Occurred Before or After the Release of the New Guidelines

Characteristic	Before New Guidelines (n=1,085)	After New Guidelines (n=721)	χ^2 *P*
Age group (y)			.007*
18–20	169 (15)	95 (13)	
21–25	408 (38)	234 (33)	
26 and older	508 (47)	392 (54)	
Race or ethnicity			.08
Hispanic	321 (30)	233 (32)	
Black	377 (35)	222 (31)	
White	307 (28)	226 (31)	
Other	80 (7)	40 (6)	
Language			.23
English	768 (71)	518 (72)	
Spanish	188 (17)	135 (19)	
Other	129 (12)	68 (9)	
Index abnormality			.11
LSIL	992 (91)	674 (93)	
HSIL	93 (9)	47 (7)	
Insurance status			.06
Uninsured	343 (32)	194 (27)	
Public	395 (36)	296 (41)	
Private	347 (32)	231 (32)	
Clinical site			<.001*
A	250 (23)	216 (30)	
B	254 (24)	161 (22)	
C	69 (6)	41 (6)	
D	118 (11)	94 (13)	
E	209 (19)	161 (22)	
F	185 (17)	48 (7)	

LSIL, low-grade squamous intraepithelial lesions; HSIL, high-grade squamous intraepithelial lesions.
Data are n (%) unless otherwise specified.
* Some observed differences may be due to differences in data collection: data collected in 2004–2005 (baseline and feasibility study) included all patients with high-grade lesions and a random selection of patients with low-grade lesions, and data from 2007–2008 (intervention) included all patients with abnormal Pap test results.

Analysis of the management of minimally abnormal Pap test results (ASC-US HPV+ and LSIL) in women younger than age 21 revealed that most health care providers adopted the new guidelines. Before guideline changes, 78% of young women with minimally abnormal Pap test results (ASC-US HPV+ and LSIL) underwent colposcopy, compared with 45% after guideline changes. Multivariable logistic regression demonstrated a 10-fold increase in the likelihood of receiving a repeated Pap test instead of colposcopy after guideline changes ($P<.001$) (see Table 2). Other factors associated with management of minimally abnormal Pap test results (ASC-US HPV+ and LSIL) in women ages 18–20 included

Table 2. Multivariable Logistic Regression on the Association of Receiving a Repeated Pap Test Compared With Colposcopy Among Women Aged 18–20 Years With Minimally Abnormal Pap Test Results (n=206*)

Characteristic	Before Guideline (n=131) n	Before Guideline (n=131) % Receiving Repeated Pap Test	After Guideline (n=75) n	After Guideline (n=75) % Receiving Repeated Pap Test	Univariable, Unadjusted OR (95% CI)	Multivariable, Adjusted† OR (95% CI)	Multivariable, Adjusted Wald χ^2 *P*
Date of abnormal Pap test finding (LSIL/ASC-US HPV+)							
Before guideline (reference)	131	22	0	0	—	—	—
After guideline	0	0	75	55	4.24 (2.30–7.84)	10.22 (4.37–23.88)	<.001
Site of care							
A	26	0	18	17	0.08 (0.02–0.27)	0.07 (0.01–0.37)	.002
B (reference)	44	30	21	90	—	—	—
C	12	33	6	100	1.29 (0.45–3.68)	1.77 (0.48–6.50)	.39
D	13	8	10	20	0.16 (0.04–0.57)	0.07 (0.01–0.33)	<.001
E	13	23	10	60	0.66 (0.25–1.75)	0.77 (0.17–3.56)	.74
F	23	35	10	50	0.67 (0.29–1.57)	1.02 (0.34–3.11)	.97
Insurance status							
Uninsured	40	23	19	53	0.80 (0.41–1.58)	1.54 (0.66–3.60)	.32
Public (reference)	62	23	40	60	—	—	—
Private	29	21	16	44	0.68 (0.32–1.46)	0.57 (0.21–1.56)	.28
Language							
English (reference)	110	26	62	55	—	—	—
Spanish	14	7	7	14	0.18 (0.04–0.80)	0.60 (0.08–4.58)	.63
Other	7	0	6	100	0.98 (0.28–3.49)	0.97 (0.24–4.01)	.97
Race or ethnicity							
Hispanic	32	13	18	28	0.28 (0.12–0.63)	0.34 (0.09–1.25)	.10
Black (reference)	66	30	29	76	—	—	—
White	26	19	24	42	0.54 (0.26–1.12)	0.64 (0.18–2.27)	.49
Other	7	0	4	100	0.72 (0.20–2.63)	1.36 (0.23–7.93)	.73

LSIL, low-grade squamous intraepithelial lesions; ASC-US, atypical squamous cells of undetermined significance; HPV, human papillomavirus.

* Forty-four patients were excluded from analysis owing to missing diagnostic data (eg, lost to follow-up), and 14 were excluded because their index Pap test indicated high-grade squamous intraepithelial lesions.

† All variables are included in multivariable analysis: date of abnormal finding, site of care, insurance status, language, and race or ethnicity.

clinical site, with patients attending Health Centers A and D being less likely to receive a repeated Pap test than those attending Health Center B. Women who spoke Spanish or were of Latina ethnicity were less likely to receive repeated Pap tests in univariable analyses but not in multivariable analyses; 50% of Latina patients attended health centers with lower guideline adherence (A and D). Assignment to the control or intervention arm of the patient navigation study was not associated with guideline adherence. The clustering analysis using generalized estimating equations controlling for clinical site of care showed similar results, with a significant increase in likelihood of receiving a Pap test instead of colposcopy after guideline change (odds ratio 6.09, 95% confidence interval 2.72–13.65). No patient characteristics were found to have an effect on receiving a Pap test before or after guideline change in the analysis using generalized estimating equations.

We performed analyses of women ages 21–25 (n=532) and 26 and older (n=702) with the goal of estimating whether guidelines had unintended effects among adult women, and we found no difference in the management of minimally abnormal Pap test results (ASC-US HPV+ and LSIL) before and after guideline changes. Among 21- to 25-year-olds, rates of colposcopy were 96% and 93% (*P*=.22) before and after guidelines; rates for women aged 26 and older were 95% and 95%, respectively (*P*=.88). Consistent with these stable, high rates of colposcopy, no demo-

graphic or clinical factors significantly predicted management with respect to date of guideline publication in multivariable analyses of these age groups.

To determine the effects of guideline changes on the management of moderate dysplasia (CIN 2) in young women, we examined the management of all cases of moderate dysplasia (CIN 2) in our dataset (n=128). Among 18–20 year olds, 6 of 16 (38%) underwent LEEPs when the initial abnormal Pap test result occurred before guideline changes, compared with 0 of 3 after guideline changes (P=.52). Among 21- to 23-year-olds, 17 of 26 (65%) underwent LEEPs when the initial abnormal Pap test result occurred before guideline change, compared with 2 of 8 (25%) after guideline changes (P=.10). No change was seen in women ages 24 and older: 31 of 44 (70%) underwent LEEPs before guideline changes compared with 23 of 31 (74%) after guideline changes (P=.72). Because women aged 18–20 and 21–23 were treated similarly, we combined these age groups in our analysis of LEEP completion. Fisher exact analysis of women aged 18–23 revealed a significant decrease in the number of LEEPs performed, from 55% before guideline changes to 18% after guideline changes (P=.04) (Fig. 1).

DISCUSSION

We used data from six community health centers serving more than 50,000 low-income and minority women to determine health care provider adherence to the 2006 American Society for Colposcopy and Cervical Pathology guidelines for conservative management of minimally abnormal Pap test results (ASC-US HPV+ and LSIL) and moderate dysplasia (CIN 2) in young women. We observed an overall decrease in the performance of colposcopies in 18- to 20-year-old women with minimally abnormal Pap test results and reduced numbers of excision procedures in 18- to 23-year-old women with moderate dysplasia after publication of the guidelines. These findings show that the new guidelines were sufficient to change clinical practices among this cohort of health care providers serving an urban population of low-income and minority women.

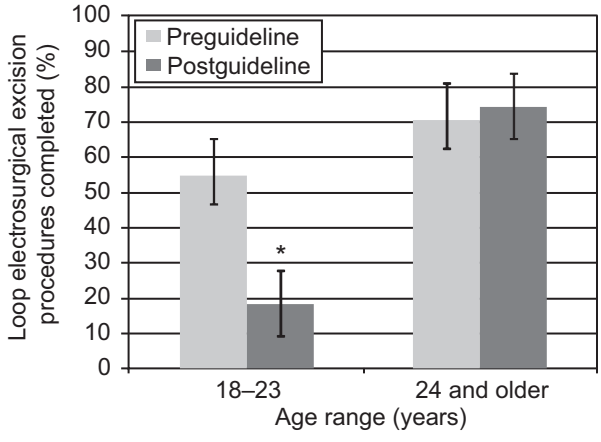


Fig. 1. The rate of loop electrosurgical excision procedures among women aged 18–23 years with moderate dysplasia (cervical intraepithelial neoplasia 2) decreased after guideline changes; no decrease was seen among older women. *P=.04.

Perkins. Adherence to Pap Test Guidelines. Obstet Gynecol 2012.

Existing literature highlights reasons why these guidelines may have been rapidly adopted. First, clinical practice guidelines that are based on clear scientific evidence are adhered to more frequently by health care providers,[19,20] and a robust literature accumulating over more than a decade indicated both high rates of regression of HPV infection and moderate dysplasia (CIN 2) in young women[5–7] as well as clear obstetric harms from treatment of cervical dysplasia with excision procedures.[9,21] These findings were highlighted in medical journals and at national meetings including those of the American Society of Colposcopy and Cervical Pathology and the American Congress of Obstetricians and Gynecologists before publication of the final consensus guidelines, thus health care providers may have been aware of the science underlying the guideline changes. In fact, our data indicated that some adolescents received repeated Pap tests for minimally abnormal Pap test results (ASC-US HPV+ and LSIL) and colposcopy for moderate dysplasia (CIN 2) before publication of guidelines. Another reason for rapid adoption of guidelines may reflect the medical mantra "first, do no harm." To avoid harming patients, many health care providers rapidly stop practices that are associated with negative outcomes, a prominent example being the dramatic decrease in use of postmenopausal hormone therapy after publication of the results of the Women's Health Initiative.[22,23] Therefore, clinicians aware of the obstetric complications associated with excision procedures (LEEP and cold knife cone) may have been eager to adopt guidelines published with the intent of avoiding these potentially harmful procedures in young women.

Multivariable analysis revealed that site of care affected the management of minimally abnormal Pap test results (ASC-US HPV+ and LSIL), with women attending Health Centers A and D being more likely to receive colposcopy instead of repeated Pap testing than were women at Health Center B. Differences in

practice have been noted at the regional, local, and individual health care provider level,[24–26] and a combination of individual and clinical systems factors may have played a role in guideline adoption in this case. Although the design of this study did not specifically examine system factors within health centers or individual health care provider factors that may have contributed to early or late adoption of guidelines, individual health care provider factors seem to be most prominent based on the workflow related to abnormal Pap test result management in these clinical sites. The six clinical sites encompassed 17 separate practices (pediatrics, family medicine, and internal medicine) in which 108 health care providers provided primary care including Pap testing during the study period. In 16 of these practices, abnormal Pap test results were routed directly back to the health care provider who ordered the Pap test, and that individual made the recommendation for colposcopy or repeated Pap test. Only one practice had a referral clinician who triaged all abnormal results. When patients with abnormal Pap test results were diagnosed with moderate dysplasia (CIN 2) via colposcopic biopsy, the decision to perform a repeated colposcopy or a LEEP was made on an individual basis by the health care provider who had taken the biopsy.

This study has several limitations related to the use of retrospective data from electronic medical records. We accessed data related to the health care providers' documentation in the patient's medical chart and procedures ordered, but we do not know the extent of health care provider–patient discussion around management options, nor the degree to which patient preference influenced therapeutic choices. We also did not collect data on patient characteristics that may have influenced the decision to perform an invasive procedure, such as a history of abnormal Pap test results, human immunodeficiency infection, or previous noncompliance with recommended follow-up. In addition, the influence of clinical site on the management of abnormal Pap test results could have reflected health center policy, health care provider education, or provider or patient attitudes, factors that could be explored in future research.

New guidelines for conservative management of minimally abnormal Pap test results and moderate cervical dysplasia in adolescents were quickly adopted among this group of community health centers serving low-income, urban populations. If this trend is observed nationwide, the resulting decrease in LEEPs may reduce racial disparities in preterm deliveries in the future.

REFERENCES

1. Kahn JA, Lan D, Kahn RS. Sociodemographic factors associated with high-risk human papillomavirus infection. Obstet Gynecol 2007;110:87–95.
2. Dunne EF, Unger ER, Sternberg M, McQuillan G, Swan DC, Patel SS, Markowitz LE. Prevalence of HPV infection among females in the United States. JAMA 2007;297:813–9.
3. SEER Stat Fact Sheet: Cervix uteri. Available at: http://seer.cancer.gov/statfacts/html/cervix.html#incidence-mortality. Retrieved October 7, 2011.
4. Clifford G, Franceschi S, Diaz M, Munoz N, Villa LL. Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases. Vaccine 2006;(24 suppl 3):S26–34.
5. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. N Engl J Med 1998;338:423–8.
6. McAllum B, Sykes PH, Sadler L, Macnab H, Simcock BJ, Mekhail AK. Is the treatment of CIN 2 always necessary in women under 25 years old? Am J Obstet Gynecol 2011;205:478.e1–7.
7. Trimble CL, Piantadosi S, Gravitt P, Ronnett B, Pizer E, Elko A, et al. Spontaneous regression of high-grade cervical dysplasia: effects of human papillomavirus type and HLA phenotype. Clin Cancer Res 2005;11:4717–23.
8. Wright TC Jr, Cox JT, Massad LS, Carlson J, Twiggs LB, Wilkinson EJ; 2001 ASCCP-sponsored Consensus Workshop. 2001 consensus guidelines for the management of women with cervical intraepithelial neoplasia. J Low Genit Tract Dis 2003;7:154–67.
9. Jakobsson M, Gissler M, Paavonen J, Tapper AM. Loop electrosurgical excision procedure and the risk for preterm birth. Obstet Gynecol 2009;114:504–10.
10. Sadler L, Saftlas A. Cervical surgery and preterm birth. J Perinat Med 2007;35:5–9.
11. Janevic T, Stein CR, Savitz DA, Kaufman JS, Mason SM, Herring AH. Neighborhood deprivation and adverse birth outcomes among diverse ethnic groups. Ann Epidemiol 2011;20:445–51.
12. Ahern J, Pickett KE, Selvin S, Abrams B. Preterm birth among African American and white women: a multilevel analysis of socioeconomic characteristics and cigarette smoking. J Epidemiol Community Health 2003;57:606–11.
13. Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D; 2006 American Society for Colposcopy and Cervical Pathology-sponsored Consensus Conference. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. Am J Obstet Gynecol 2007;197:346–55.
14. Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D; 2006 American Society for Colposcopy and Cervical Pathology-sponsored Consensus Conference. 2006 consensus guidelines for the management of women with cervical intraepithelial neoplasia or adenocarcinoma in situ. J Low Genit Tract Dis 2007;11:223–39.
15. Meissner HI, Tiro JA, Yabroff KR, Haggstrom DA, Coughlin SS. Too much of a good thing? Physician practices and patient willingness for less frequent pap test screening intervals. Med Care 2010;48:249–59.
16. Saraiya M, Berkowitz Z, Yabroff KR, Wideroff L, Kobrin S, Benard V. Cervical cancer screening with both human papillomavirus and Papanicolaou testing vs Papanicolaou testing alone: what screening intervals are physicians recommending? Arch Intern Med 2010;170:977–85.
17. Freund KM, Battaglia TA, Calhoun E, Dudley DJ, Fiscella K, Paskett E, et al; Patient Navigation Research Program Group.

National Cancer Institute Patient Navigation Research Program: methods, protocol, and measures. Cancer 2008;113:3391–9.

18. Battaglia TA, Santana MC, Bak S, Gokhale M, Lash TL, Ash AS, et al. Predictors of timely follow-up after abnormal cancer screening among women seeking care at urban community health centers. Cancer 2010;116:913–21.
19. Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 1998;317:858–61.
20. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282:1458–65.
21. Sadler L, Saftlas A, Wang W, Exeter M, Whittaker J, McCowan L. Treatment for cervical intraepithelial neoplasia and risk of preterm delivery. JAMA 2004;291:2100–6.
22. Haas JS, Kaplan CP, Gerstenberger EP, Kerlikowske K. Changes in the use of postmenopausal hormone therapy after the publication of clinical trial results. Ann Intern Med 2004;140:184–8.
23. Hersh AL, Stefanick ML, Stafford RS. National use of postmenopausal hormone therapy: annual trends and response to recent evidence. JAMA 2004;291:47–53.
24. Stafford RS, Blumenthal D, Pasternak RC. Variations in cholesterol management practices of U.S. physicians. J Am Coll Cardiol 1997;29:139–46.
25. Goyert GL, Bottoms SF, Treadwell MC, Nehra PC. The physician factor in cesarean birth rates. N Engl J Med 1989;320:706–9.
26. Strasser DC, Smits SJ, Falconer JA, Herrin JS, Bowen SE. The influence of hospital culture on rehabilitation team functioning in VA hospitals. J Rehabil Res Dev 2002;39:115–25.

Parental Intermittent Claudication as Risk Factor for Claudication in Adults

Scott G. Prushik, MD[a,†], Alik Farber, MD[b,†], Philimon Gona, PhD[c,d], Peter Shrader, BA[c,e], Michael J. Pencina, PhD[c,d], Ralph B. D'Agostino, Sr, PhD[c,d], and Joanne M. Murabito, MD, ScM[d,f,*]

Little is known about the familial aggregation of intermittent claudication (IC). Our objective was to examine whether parental IC increased the risk of IC in adult offspring, independent of the established cardiovascular risk factors. We evaluated the Offspring Cohort Participants of the Framingham Heart Study who were ≥30 years old, cardiovascular disease free, and had both parents enrolled in the Framingham Heart Study (n = 2,970 unique participants, 53% women). Pooled proportional hazards regression analysis was used to examine whether the 12-year risk of incident IC in offspring participants was associated with parental IC, adjusting for age, gender, diabetes, smoking, systolic blood pressure, total cholesterol, high-density lipoprotein cholesterol, and antihypertensive and lipid treatment. Of the 909 person-examinations in the parental IC history group and 5,397 person-examinations in the no-parental IC history group, there were 101 incident IC events (29 with parental IC history and 72 without a parental IC history) during follow-up. The age- and gender-adjusted 12-year cumulative incidence rate per 1,000 person-years was 5.08 (95% confidence interval [CI] 2.74 to 7.33) and 2.34 (95% CI 1.46 to 3.19) in participants with and without a parental IC history. A parental history of IC significantly increased the risk of incident IC in the offspring (multivariable adjusted hazard ratio 1.81, 95% CI 1.14 to 2.88). The hazard ratio was unchanged, with an adjustment for the occurrence of cardiovascular disease (hazard ratio 1.83, 95% CI 1.15 to 2.91). In conclusion, IC in parents increases the risk of IC in adult offspring, independent of the established risk factors. These data suggest a genetic component of peripheral artery disease and support future research into genetic causes. **(Am J Cardiol 2012;109:736–741)**

A National Institutes of Health consensus statement has described family history as vital to patient care, because it can uncover information about factors that contribute to the risk of developing common diseases, such as diabetes mellitus, stroke, cancer, and heart disease.[1] The occurrence of cardiovascular disease (CVD) in a parent or sibling confers an increased risk of CVD in middle-age adults, distinct from the traditional risk factors.[2,3] The parental occurrence of stroke is associated with a threefold increase in the risk of offspring stroke.[4] However, little is known about the familial aggregation of peripheral artery disease (PAD). The subjects whose siblings were diagnosed with premature PAD were shown to have an almost threefold increase in PAD.[5] Finally, the early onset of symptomatic CVD is more common in first-degree relatives of patients with premature PAD than in the relatives of healthy persons.[6] The latter 2 studies were limited by the small sample size, an examination of PAD prevalence rather than incidence, and an inability to quantify the degree of familial aggregation of PAD that was independent of the established risk factors. The Framingham Heart Study (FHS) affords the unique opportunity to study intermittent claudication (IC) across 2 generations using prospectively collected data in a large community-based sample. The present study was undertaken to test the hypothesis that parental IC confers an increased risk of IC in adult offspring, independent of established CVD risk factors.

Methods

The FHS is a prospective epidemiologic cohort study that was established in 1948 when 5,209 residents of Framingham, Massachusetts, aged 28 to 62 years, were enrolled. The members of the original cohort have undergone examinations every 2 years. In 1971, the offspring of the original cohort (n = 3,548) and the spouses of the offspring (n = 1,576), aged 5 to 70 years, were enrolled into the Framingham Offspring Study.[7] The offspring cohort has undergone an examination about every 4 to 8 years. All participants provided informed consent, and the institutional

[a]Department of Vascular Surgery, University of Massachusetts, Worchester, Massachusetts; Departments of [b]Vascular Surgery and [f]Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, Massachusetts; [c]Department of Mathematics and Statistics, Boston University, Boston, Massachusetts; [d]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts; and [e]New England Research Institutes, Watertown, Massachusetts. Manuscript received September 14, 2011; revised manuscript received and accepted October 25, 2011.

This work was supported by grant N01-HC-25195 from the National Heart, Lung, and Blood Institute's Framingham Heart Study (Bethesda, Maryland).

*Corresponding author: Tel: (508) 935-3461; fax: (508) 626-1262.
E-mail address: murabito@bu.edu (J.M. Murabito).

† Drs. Prushik and Farber contributed equally to this article.

0002-9149/12/$ – see front matter
doi:10.1016/j.amjcard.2011.10.032

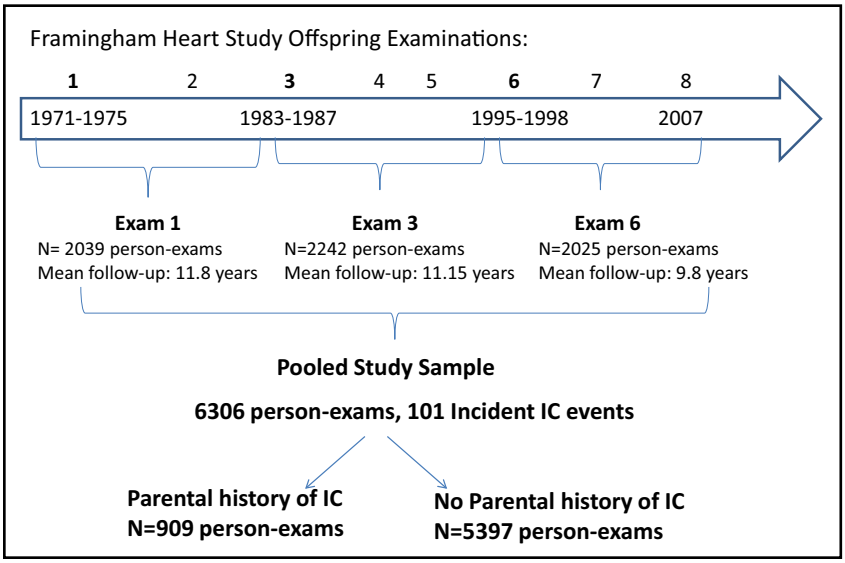


Figure 1. Study sample. Data from 3 baseline examinations, each with 12 years of follow-up were pooled.

review board of Boston University School of Medicine approved all study protocols.

The data from 3 offspring examinations (Figure 1), each with 12 years of follow-up, were pooled: offspring examination 1 (1971 to 1975), 3 (1983 to 1987), and 6 (1995 to 1998). The follow-up for the final examination ended in December 2007. Given the structure of the follow-up examinations in the offspring, we chose to study the 12-year incidence of IC to have a comparable length of nonoverlapping follow-up after each baseline examination. All offspring participants who were ≥30 years old at any of the 3 baseline examinations were eligible if both parents were enrolled in the original cohort and if the Offspring participant was free of CVD and IC at the examination. A total of 2,970 unique subjects (1,405 men) were included in our final study sample, including 346 with a parental history of IC and 2,624 with both parents free of IC. Parental IC was defined as the occurrence of IC in a parent before the offspring examination. Both parental IC events and incident IC events occurring in the offspring were adjudicated by a panel of 3 senior investigators using the same previously established criteria. The investigators were unaware of the parental IC status. All available information was used to determine the presence of IC, including the standardized physician-administered IC questionnaire that was a part of each routine FHS research clinic visit, any available records from the office visits with the participants' personal healthcare provider, and the hospital records pertaining to PAD.[8] The standardized physician-administered questionnaire asked about the presence of calf and leg discomfort brought on by exertion, the relation of the discomfort to the rapidity of walking or uphill walking, and whether the symptoms were relieved with rest. The final diagnosis of IC was determined from the clinical history only, without confirmatory testing. At the more contemporary FHS examinations, the participants were queried about lower extremity revascularization procedures and all self-reports were validated from the medical records.

At each offspring examination, the risk factors were directly measured and the occurrence of CVD was updated. Blood pressure at rest was measured twice. Current smoking was defined as smoking ≥1 cigarettes/day in the year preceding the examination. Blood was drawn, with the patient in the fasting state, for total cholesterol, high-density lipoprotein cholesterol, and triglyceride measurements. Diabetes was defined as a fasting glucose level of ≥126 mg/dl (7.0 mmol/L) or the use of insulin or oral hypoglycemic agents. CVD was defined as any of the following events: myocardial infarction, coronary insufficiency, angina pectoris, stroke, transient ischemic attack, congestive heart failure, or cardiovascular death.

The follow-up time within each 12-year period was calculated as the interval from each baseline visit date until the diagnosis date of IC for those participants who developed the disease and censored at the earliest of the date of the last examination, date of death, or end of the 12-year period for participants who did not develop the disease. The age- and gender-adjusted incidence rates and 95% confidence intervals (CIs) per 1,000 person-years were calculated in each parental IC history group by dividing the number of IC events observed by the total person-years. The Kaplan-Meier curves and the log-rank test were used to plot and compare the cumulative incidence rates. Pooled proportional hazards regression analyses were used to examine whether the 12-year risk of incident IC in offspring was associated with parental IC. This method of pooling person-examinations provides estimates of the effect similar to a time-dependent Cox proportional hazards model.[9] Furthermore, this method allowed us to update the risk factors and parental IC at each examination. The hazard ratios (HRs) and 95% CIs were calculated with the reference group consisting of participants with no parental IC before the examination. The covariates used in the multivariable model included age, gender, diabetes, current smoking, systolic blood pressure, antihypertensive treatment, total cholesterol, high-density lipoprotein cholesterol, and cholesterol lowering treatment. In the secondary analyses, we adjusted for the occurrence of CVD in the Offspring par-

Table 1
Baseline characteristics by parental history of intermittent claudication (IC)

Variable	Parental History of IC*		p Value
	No (n = 5,397 Person-Examinations)	Yes (n = 909 Person-Examinations)	
Age (years)	47.6 ± 11	49.9 ± 10.5	<0.0001
Women	2,927 (54%)	466 (51%)	0.10
Current smoker	1,501 (28%)	276 (30%)	0.12
Diabetes mellitus	202 (4%)	62 (7%)	<0.0001
Systolic blood pressure (mm Hg)	123 ± 17	126 ± 17	<0.0001
Diastolic blood pressure (mm Hg)	78 ± 10	79 ± 10	0.003
Total cholesterol (mg/dl)	206 ± 40	210 ± 41	0.005
High-density lipoprotein cholesterol (mg/dl)	52 ± 15	50 ± 16	0.007
Triglycerides (mg/dl)	119 ± 126	131 ± 105	0.007
Lipid-lowering medications	139 (3.6%)	42 (4.6%)	0.001
Body mass index (kg/m^2)	26 ± 4.8	27 ± 4.7	0.007
Antihypertensive medications	634 (12%)	162 (18%)	<0.0001

* A total of 346 unique subjects were included in parental history group and 2,624 unique subjects were in no-parental history group.

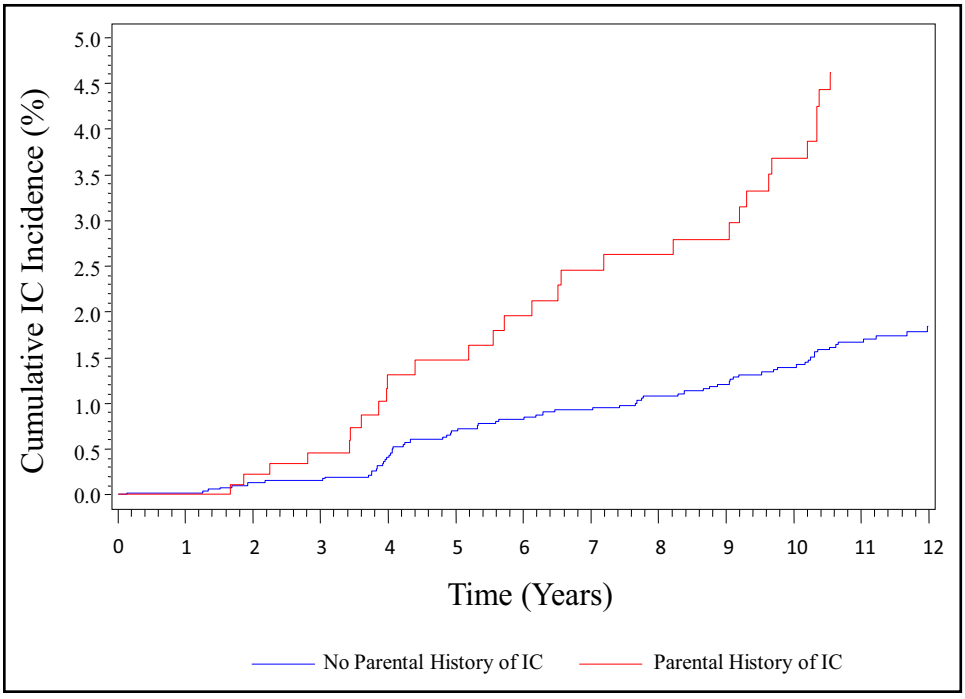


Figure 2. Age- and gender-adjusted cumulative incidence per 1,000 person-years of intermittent claudication.

ticipants using 2 approaches. First, we used Cox models in which CVD was entered as a time-dependent covariate. Next, the follow-up time was censored when an Offspring participant developed any CVD event to account for the fact that CVD increases the risk of IC. To assess the incremental predictive utility of parental IC history associated with incident IC in the offspring, we calculated the c-statistic for the model with clinical covariates alone and the full model with clinical covariates and parental IC history.[10,11] We assessed the model calibration (i.e., concordance of observed risk and that predicted by the model with parental IC history) by calculating the Hosmer-Lemeshow Chi-squared statistic for the Cox models.[10,11] To evaluate whether the inclusion of parental IC history improved the risk classification of participants, we calculated the enhanced "net reclassification improvement" (NRI) using an extension to survival analysis that uses Kaplan-Meier estimates of event probabilities at 12 years.[11] We used 12-year IC risk thresholds of <2%, 2% to 5%, and >5% for the NRI index. The NRI is used to assess how well a new marker "reclassifies" patients from 1 risk category to another. Because no categories for the absolute risk of IC have been previously established, we also assessed the "category-less" NRI, which assesses any upward or downward reclassification. Values >0 correspond to improved reclassification.[12] We performed a secondary analysis defining the incident events in the Offspring as IC and or lower extremity revascularization. All statistical analyses were performed using SAS statistical software, version 9 (SAS

Table 2
Twelve-year risk of intermittent claudication (IC) by parental occurrence of intermittent claudication (IC)

Model	HR	95% CI	p Value
Age and gender adjusted	2.29	1.49–3.54	0.0002
Multivariable adjusted*	1.81	1.14–2.88	0.01
Multivariable plus cardiovascular disease (time-dependent)	1.83	1.15–2.91	0.01
Multivariable plus censor for interim cardiovascular disease	1.92	1.16–3.19	0.01

* Multivariable model adjusted for following covariates: age, gender, diabetes, current cigarette smoking, systolic blood pressure, antihypertensive treatment, total cholesterol, high-density lipoprotein cholesterol, and lipid-lowering treatment.

Institute, Cary, North Carolina). Statistical significance was defined as a 2-tailed p value <0.05.

Results

The baseline characteristics of the offspring study sample are listed in Table 1. The group of participants with parental history of IC was older (mean age 49.9 vs 47.6 years, $p<0.0001$) and, with the exception of current smoking, had significantly greater risk factor levels compared to the group of participants without a parental history of IC. During the follow-up period, 101 incident IC events (29 in participants with a parental history of IC and 72 in participants with no parental history of IC) occurred. The age- and gender-adjusted 12-year cumulative incidence rate per 1,000 person-years was 5.08 (95% CI 2.74 to 7.33) in participants with a parental history of IC and 2.34 (95% CI 1.46 to 3.19) in participants with no parental history of IC (log-rank test, $p<0.0001$; Figure 2). Parental IC was associated with a significantly increased risk of incident IC in offspring (age- and gender-adjusted HR 2.29). The association was modestly attenuated but remained significant after adjustment for traditional risk factors (multivariable-adjusted HR 1.81, 95% CI 1.14 to 2.88; Table 2). The association was unchanged after additional adjustment for interim CVD and the magnitude and significance of the effect of parental IC persisted in the analysis in which the follow-up time was censored when the offspring participant developed an incident CVD event (Table 2). The addition of parental IC history to a multivariable model incorporating the baseline covariates increased the already high c-statistic from 0.831 (95% CI 0.794 to 0.868) to 0.837 (95% CI 0.801 to 0.873). The 2 c-statistics were not significantly different statistically ($p = 0.22$). The model with parental IC history had excellent calibration (Hosmer-Lemeshow Chi-squared 14.07; $p = 0.20$). The category-based NRI was modest (9.3%, 95% CI 1.9% to 17.3%) and the category-free NRI was 34.5% (95% CI 15.5% to 55.6%). The NRI estimates remained essentially unchanged, with adjustment for the occurrence of CVD entered as a time-dependent covariate or in a model in which the follow-up time was censored when a participant developed any CVD event. The association between parental IC and incident PAD, defined as IC and/or lower extremity revascularization ($n = 114$ events), was very similar to the primary analysis (multivariable-adjusted HR 1.76, 95% CI 1.14 to 2.72; additional adjustment for CVD as a time-dependent covariate [HR 1.77, 95% CI 1.15 to 2.73] and censoring at the first occurrence of CVD [HR 1.84, 95% CI 1.13 to 2.97] did not substantively change the association).

Discussion

In our community-based sample, parental IC confers a nearly twofold increased risk of IC in adult offspring even after accounting for traditional risk factors and the interim occurrence of CVD. Parental history of IC is quite powerful in predicting IC in adult offspring; however, the small improvement in the c-statistic and category-based NRI suggested a modestly improved prediction of offspring IC beyond that of the established risk factors. Our results suggest the presence of a genetic predisposition to IC. However, in addition to the genetic factors, environmental and lifestyle factors shared within families can contribute to the susceptibility to common diseases.[13] A person's family history provides a readily accessible and important clinical tool to gain insight into a person's risk of developing disease, including PAD. Knowledge of one's family history has the potential to motivate positive lifestyle changes, enhance individual empowerment, and influence clinical decisions with respect to preventative interventions.[1] This might be especially important for PAD, a disease that is often undetected and undertreated by clinicians[14] and underappreciated by the lay public for its association with the risk of myocardial infarction, stroke, and death.[15]

Little is known concerning the genetic determinants of PAD, but family aggregation and heritability estimates suggest a significant genetic contribution.[16,17] Recently, Zintzaras and Zdoukopoulos[18] created a publically available database that catalogs the genetic association studies of PAD. However, to date, the genetic associations for PAD remain inconclusive, with no strongly replicated results.[19] The genetic susceptibility to PAD and CVD is likely conferred by multiple genes interacting with a variety of environmental factors. Although PAD might share some genetic variants with CVD, likely a set of genetic variants exists that are unique to PAD, given the phenotypic diversity across the vascular system.[19]

Genome-wide association studies have identified a strongly replicated association between variants on chromosome 9p21 and myocardial infarction and coronary artery disease.[20] The chromosome 9p21 locus appears to play a broad role in arterial disease, because it is also associated with atherosclerotic stroke,[21] vascular stiffness,[22] and cerebral and abdominal aortic aneurysm.[23,24] The chromosome 9p21 locus was associated with clinically diagnosed PAD; however, the relation was no longer present when the coronary artery disease cases were removed from the sample.[24] Recently, this locus was reported to be associated with a greater prevalence of PAD and a lower mean arterial brachial index level in older adults from 3 population-based samples that was independent of risk factors and persisted even after removal of those with myocardial infarction.[25] The genome-wide association approach successfully identified genetic variants on chromosome 15q24 for nicotine dependence that conferred the risk of PAD.[26] Thus, the

genetic variants identified appeared to influence the risk of PAD through shared biologic mechanisms for CVD and smoking-related behaviors. Additionally, in a recent study of 1,292 patients with abdominal aortic aneurysms, the A allele of rs7025486 on chromosome 9q33 was associated with abdominal aortic aneurysm and PAD.[27] No association was found between this gene and the CVD risk factors, suggesting an independent genetic predisposition exists.[27]

Using a murine hind limb ischemia model of PAD, Dokun et al[28] identified genetic influences on PAD disease severity. Additional work is needed to elucidate the genes responsible for the association and their human orthologs. One study, using microarray analysis of femoral artery specimens, elucidated >400 genes that are either up- or down-regulated in patients with severe atherosclerosis.[29] Another similar study demonstrated that many of the genes involved in atherosclerosis are also involved in lipid synthesis and immune function pathways.[30] A clinical investigation that leads to the elucidation of the genetic basis of PAD might provide potential targets for future focused gene therapies.

Several strengths and limitations of the present study merit comment. The FHS is a large, community-based sample, with the data collected prospectively from both parents and offspring. The symptoms of IC were adjudicated by a panel of senior investigators using well-established criteria, and the risk factors were directly measured, rather than obtained by self-report from the participants, which could be subject to misclassification. Most FHS participants were of white, European ancestry, limiting the generalizability of the results to those from other race/ethnic backgrounds. The diagnosis of IC is determined by symptoms alone, without confirmatory testing; thus, some of the participants with IC in the present study might not have had PAD. In contrast, only about 50% of those with PAD have leg pain symptoms. The risk factors used for adjustment in the multivariable models were from a single-occasion measurement and might not represent an assessment of lifetime risk factor exposure. Hence, it is possible that the effect of a parental history of IC on the personal risk of IC might have been overestimated. Residual confounding might be present in that the parental history might reflect a low social class, passive smoking, diet, or sedentary behaviors that we were not able to account for in our study.

1. Berg AO, Baird MA, Botkin JR, Driscoll DA, Fishman PA, Guarino PD, Hiatt RA, Jarvik GP, Millon-Underwood S, Morgan TM, Mulvihill JJ, Pollin TI, Schimmel SR, Stefanek ME, Vollmer WM, Williams JK. National Institutes of Health State-of-the-Science Conference Statement: family history and Improving Health. *Ann Intern Med* 2009;151:872–877.
2. Lloyd-Jones DM, Nam BH, D'Agostino RB Sr, Levy D, Murabito JM, Wang TJ, Wilson PW, O'Donnell CJ. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* 2004;291:2204–2211.
3. Murabito JM, Pencina MJ, Nam BH, D'Agostino RB Sr, Wang TJ, Lloyd-Jones D, Wilson PW, O'Donnell CJ. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA* 2005;294:3117–3123.
4. Seshadri S, Beiser A, Pikula A, Himali JJ, Kelly-Hayes M, Debette S, DeStefano AL, Romero JR, Kase CS, Wolf PA. Parental occurrence of stroke and risk of stroke in their children: the Framingham study. *Circulation* 2010;121:1304–1312.
5. Valentine RJ, Guerra R, Stephan P, Scoggins E, Clagett GP, Cohen J. Family history is a major determinant of subclinical peripheral arterial disease in young adults. *J Vasc Surg* 2004;39:351–356.
6. Valentine RJ, Verstraete R, Clagett GP, Cohen JC. Premature cardiovascular disease is common in relatives of patients with premature peripheral atherosclerosis. *Arch Intern Med* 2000;160:1343–1348.
7. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families: the Framingham Offspring study. *Am J Epidemiol* 1979;110:281–290.
8. Murabito JM, D'Agostino RB, Silbershatz H, Wilson WF. Intermittent claudication: a risk profile from the Framingham Heart Study. *Circulation* 1997;96:44–49.
9. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med* 1990;9:1501–1515.
10. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;23:2109–2123.
11. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–172.
12. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;30:11–21.
13. Cupples LA. Family study designs in the age of genome-wide association studies: experience from the Framingham Heart Study. *Curr Opin Lipidol* 2008;19:144–150.
14. Hirsch AT, Criqui MH, Treat-Jacobson D, Regensteiner JG, Creager MA, Olin JW, Krook SH, Hunninghake DB, Comerota AJ, Walsh ME, McDermott MM, Hiatt WR. Peripheral arterial disease detection, awareness, and treatment in primary care. *JAMA* 2001;286:1317–1324.
15. Hirsch AT, Murphy TP, Lovell MB, Twillman G, Treat-Jacobson D, Mohler ER, Creager MA, Hobson RW, Robertson RM, Howard WJ, Schroeder P, Criqui MH, Criqui MH; Peripheral Arterial Disease Coalition. Gaps in public knowledge of peripheral arterial disease: the first national PAD public awareness survey. *Circulation* 2007;116:2086–2094.
16. Kullo IJ, Turner ST, Kardia SL, Mosley TH Jr, Boerwinkle E, de Andrade M. A genome-wide linkage scan for ankle-brachial index in African American and non-Hispanic white subjects participating in the GENOA study. *Atherosclerosis* 2006;187:433–438.
17. Murabito JM, Guo CY, Fox CS, D'Agostino RB. Heritability of the ankle-brachial index: the Framingham Offspring study. *Am J Epidemiol* 2006;164:963–968.
18. Zintzaras E, Zdoukopoulos N. A field synopsis and meta-analysis of genetic association studies in peripheral arterial disease: the CUMAGAS-PAD database. *Am J Epidemiol* 2009;170:1–11.
19. Knowles JW, Assimes TL, Li J, Quertermous T, Cooke JP. Genetic susceptibility to peripheral arterial disease: a dark corner in vascular biology. *Arterioscler Thromb Vasc Biol* 2007;27:2068–2078.
20. Schunkert H, Götz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, EL Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, König IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ; Cardiogenics Consortium. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675–1684.
21. Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, Meitinger T, Wichmann E, Mitchell BD, Furie K, Slowik A, Rich SS, Syme PD, MacLeod MJ, Meschia JF, Rosand J, Kittner SJ, Markus HS, Müller-Myhsok B, Dichgans M; International Stroke Genetics Consortium. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. *Ann Neurol* 2009;65:531–539.
22. Björck HM, Länne T, Alehagen U, Persson K, Rundkvist L, Hamsten A, Dahlström U, Eriksson P. Association of genetic variation on chromosome 9p21.3 and arterial stiffness. *J Intern Med* 2009;265:373–381.

23. Thompson AR, Golledge J, Cooper JA, Hafez H, Norman PE, Humphries SE. Sequence variant on 9p21 is associated with the presence of abdominal aortic aneurysm disease but does not have an impact on aneurysmal expansion. *Eur J Hum Genet* 2009;17:391–394.
24. Helgadottir A, Thorleifsson G, Magnusson KP, Grétarsdottir S, Steinthorsdottir V, Manolescu A, Jones GT, Rinkel GJ, Blankensteijn JD, Ronkainen A, Jaakelainen JE, Kyo Y, Lenk GM, Sakalihasan N, Kostulas K, Gottsäter A, Flex A, Stefansson H, Hansen T, Andersen G, Weinsheimer S, Borch-Johnsen K, Jorgensen T, Shah SH, Quyyumi AA, Granger CB, Reilly MP, Austin H, Levey AI, Vaccarino V, Palsdottir E, Walters GB, Jonsdottir T, Snorradottir S, Magnusdottir D, Gudmundsson G, Ferrell RE, Sveinbjornsdottir S, Hernesniemi J, Niemelä M, Limet R, Andersen K, Sigurdsson G, Benediktsson R, Verhoeven EL, Tejink JA, Grobbee DE, Rader DJ, Collier DA, Pedersen O, Pola R, Hillert J, Lindblad B, Valdimarsson EM, Magnadottir HB, Wijmenga C, Tromp G, Baas AF, Ruigrok YM, van Rij AM, Kuivaniemi H, Powell JT, Matthiasson SE, Gulcher JR, Thorgeirsson G, Kong A, Thorsteinsdottir U, Stefansson K. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* 2008;40:217–224.
25. Cluett C, McDermott MM, Guralnik J, Ferrucci L, Bandinelli S, Miljkovic I, Zmuda JM, Li R, Tranah G, Harris T, Rice N, Henley W, Frayling TM, Murray A, Melzer D. The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people. *Circ Cardiovasc Genet* 2009;2:347–353.
26. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsäter A, Flex A, Aben KK, de Vegt F, Melders PF, Isla D, Vidal MJ, Asin L, Saez B, Murilla L, Blondal T, Kolbeinsson H, Stefansson JG, Hansdottir I, Runarsdottir V, Pola R, Lindblad B, van Rij Am DB, Haltmayer M, Mayordomo JI, Kiemeney LA, Matthiasson SE, Oskarsson H, Tyrfingsson T, Gudbjartsson DF, Gulcher JR, Jonsson S, Thorsteinsdottir U, Kong A, Stefansson K, Stefansson K. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 2008;452:638–642.
27. Gretarsdottir S, Baas AF, Thorleifsson G, Holm H, den Heijer M, de Vries JP, Kranendonk SE, Zeebregts CJ, van Sterkenburg SM, Geelkerken RH, van Rij AM, Williams MJ, Boll AP, Kostic JP, Jonasdottir A, Jonasdottir A, Walters GB, Masson G, Sulem P, Saemundsdottir J, Mouy M, Magnusson KP, Tromp G, Elmore JR, Sakalihasan N, Limet R, Defraigne JO, Ferrell RE, Ronkainen A, Ruigrok YM, Wijmenga C, Grobbee DE, Shah SH, Granger CB, Quyyumi AA, Vaccarino V, Patel RS, Zafari AM, Levey AI, Austin H, Girelli D, Pignatti PF, Oliveri O, Martinelli N, Malerba G, Trabetti E, Becker LC, Becker DM, Reilly MP, Rader DJ, Mueller T, Dieplinger B, Haltmayer M, Urbonavicius S, Lindblad B, Gottsater A, Gaetani E, Pola R, Wells P, Roger M, Forgie M, Langlois N, Corral J, Vicente V, Fontcuberta J, Espana F, Grarup N, Jorgensen T, Witte DR, Hansen T, Pedersen O, Aben KK, de Graaf J, Holewijn S, Folkersen L, Franco-Cereceda A, Eriksson P, Collier DA, Stefansson H, Steinthorsdottir V, Rafnar T, Valdimarsson EM, Magnadottir HB, Sveinbjornsdottir S, Olafsson I, Magnusson MK, Palmason R, Haraldsdottir V, Andersen K, Onundarson PT, Thorgeirsson G, Kiemeney LA, Powell JT, Carey DJ, Kuivaniemi H, Lindholt JS, Jone GT, Kong A, Blankensteijn JD, Matthiasson SE, Thorsteinsdottir U, Stefansson K. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. *Nat Genet* 2010;42:692–697.
28. Dokun AO, Keum S, Hazarika S, Li Y, Lamonte GM, Wheeler F, Marchuk DA, Annex BH. A quantitative trait locus (LSq-1) on mouse chromosome 7 is linked to the absence of tissue loss after surgical hindlimb ischemia. *Circulation* 2008;117:1207–1215.
29. Fu S, Zhao H, Shi J, Abzhanov A, Crawford K, Ohno-Machado L, Zhou J, Du Y, Kuo WP, Zhang J, Jiang M, Jin JG. Peripheral arterial occlusive disease: global gene expression analyses suggest a major role for immune and inflammatory responses. *BMC Genomics* 2008;9:369.
30. Evans DC, Sileshi B, Zakaria AM, Giangiacomo D, Manson RJ, Lawson JH. Genomic modeling of atherosclerosis in peripheral arterial disease and its variant phenotype in patients with diabetes. *Vascular* 2008;16:225–235.

AIDS Care
2012, 1–6, iFirst

Associations between partner violence perpetration and history of STI among HIV-infected substance using men in Russia

Anita Raj[a,b]*, Jeremy D. Kidd[b,c], Debbie M. Cheng[d], Sharon Coleman[e], Carly Bridden[b], Elena A. Blokhina[f], Evgeny Krupitsky[g] and Jeffrey H. Samet[b,h]

[a]*Division of Global Public Health, Department of Medicine, University of California – San Diego School of Medicine, San Diego, CA, USA;* [b]*Boston University School of Medicine/Boston Medical Center, Department of Medicine, Section of General Internal Medicine, Clinical Addiction Research and Education (CARE) Unit, Boston, MA, USA;* [c]*School of Medicine, Virginia Commonwealth University, Richmond, VA, USA;* [d]*Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA;* [e]*Data Coordinating Center, Boston University School of Public Health, Boston, MA, USA;* [f]*Institute of Pharmacology, Pavlov Medical University, St. Petersburg, Russia;* [g]*Pavlov Medical University, St. Petersburg, Russia;* [h]*Department of Community Health Sciences, Boston University School of Public Health, Boston, MA, USA*

(*Received 18 January 2012; final version received 14 August 2012*)

Studies document a significant association between victimization from intimate partner violence (IPV) and sexually transmitted infections (STIs) and HIV among substance using women in Russia and elsewhere, but no study has examined IPV perpetration and STI among Russian men or HIV-infected men in Eastern Europe. This study was designed to assess the association between lifetime history of IPV perpetration and STI (lifetime and current) among substance using HIV-infected men in Russia. Cross-sectional analyses were conducted with baseline data from 415 male participants enrolled in a randomized HIV intervention clinical trial [the HERMITAGE Study]. Participants were HIV-infected men reporting recent heavy alcohol use and unprotected sex in St. Petersburg, Russia. Baseline surveys assessed demographics, IPV perpetration, risk behaviors, and STI history. Current STI was assessed via blood testing for syphilis and urine testing for gonorrhea, Chlamydia and Trichomonas. Multiple logistic regression analyses were used to assess the association between history of IPV with lifetime and current STI. Participants were aged 20–57 years. Almost half of participants (46%) reported a history of IPV perpetration; 81% reported past 30-day binge alcohol use, and 43% reported past 30-day injection drug use. Past and current STI was 41% and 12%, respectively. Men reporting a history of IPV perpetration had significantly higher odds of reporting ever having an STI (AOR = 1.6, 95% CI = 1.1, 2.4) but lower odds of testing positive for a current STI (AOR = 0.50, 95% CI = 0.26, 0.96). These findings demonstrate that a history of male IPV perpetration is common in HIV-infected Russian men and associated with a history of STI. Programmatic work toward IPV prevention is needed in Russia and may be beneficial in mitigating STIs, but more research is needed to understand how and why the association between IPV and STI changes over time in this population.

Keywords: HIV; Russia; IPV; substance use; STI

Introduction

More than one-third of women in Russia have been victims of male partner-perpetrated intimate partner violence (IPV) (Horne, 1999; Kalichman, Kelly, Shaboltas, & Granskaya, 2000; Lokhmatkina, Kuznetsova, & Feder, 2010), and victims of IPV are at increased risk for contracting HIV (World Health Organization [WHO], 2005). Mechanisms for the association between IPV and HIV include both greater sexual risk taking (e.g., sex trade involvement, non-use of condoms) and substance use (e.g., injection drug use, risky sex while intoxicated) among victims of IPV, according to findings from Russia (Kalichman et al., 2000; Zhan et al., 2012), Ukraine (Dude, 2007), South Africa (Jewkes, Dunkle, Nduna, et al., 2006), and the USA (Bogart et al., 2005; Dude, 2007; El-Bassel, Witte, Wada, Gilbert, & Wallace, 2001; El-Bassel, Gilbert, Wu, Go, & Hill, 2005a; Panchanadeswaran et al., 2010; Schmuel & Schenker, 1998; Silverman et al., 2011; Silverman, Raj, Mucci, & Hathaway, 2001). However, research from Russia and elsewhere also indicate that male perpetrators of IPV are more likely to engage in risky sex (e.g., concurrent sex partnering, non-use of condoms, buying sex) and substance use (Decker et al., 2008, 2009; Dunkle et al., 2006; El-Bassel, Fontdevila, et al., 2001; Jewkes, Dunkle, Koss, et al., 2006; Raj et al., 2006; Raj, Reed, Welles, Santana, & Silverman, 2008; Santana, Raj, Decker, La Marche, & Silverman, 2006; Silverman, Decker, Kapur, Gupta, & Raj, 2007; Zhan et al., 2012). Longitudinal studies with women and men in the US further indicate the bidirectional nature of these associations

*Corresponding author. Email: anitaraj@ucsd.edu

ISSN 0954-0121 print/ISSN 1360-0451 online

http://dx.doi.org/10.1080/09540121.2012.722188
http://www.tandfonline.com

(El-Bassel, Gilbert, Wu, Go, & Hill, 2005b; Gilbert, El-Bassel, Wu, & Chang, 2007). Little of this research has focused on HIV-infected populations outside the US and in Eastern Europe, and may have growing importance given Russia's increasingly heterosexually driven epidemic (Goliusov et al., 2008). Thus, this study assesses lifetime IPV perpetration and its association with STI and high risk sexual behaviors among a sample of male HIV-infected substance users in St. Petersburg, Russia.

Methods

This study involved a cross-sectional analysis of baseline data from male participants of a randomized clinical trial of an HIV risk reduction intervention for HIV-infected risky drinkers in Russia (the HERMITAGE Study). Between July 2007 and April 2010, eligible adults were enrolled from four HIV or substance use clinical sites and one needle exchange program in St. Petersburg, Russia. Trained research associates recruited participants; those indicating interest were escorted into a private room to assess eligibility and to obtain informed consent. Subsequent to consent, trained staff conducted the baseline survey assessment; urine and blood samples were then collected for STI testing. Baseline incentives of 300 rubles (US$7) were provided. Of 921 individuals screened, 732 were eligible, and 700 agreed to participate. Analyses involved data from men ($n = 415$), all men reported female sex partners. Institutional Review Boards of Boston University Medical Campus and Pavlov State Medical University approved this study.

Measures

Single items assessed subjects' age, marital status, education, incarceration, employment status, and whether the participant was currently on antiretroviral therapy (ART). Past 30-day drinking and binge drinking (5+ drinks per day) were assessed using the 30-day Timeline Follow Back (Sobell & Sobell, 1992, 1996). Past 30-day injection drug use (IDU) was assessed using an item from the Risk Behavior Survey (RBS) (Petry, 2001; Weatherby et al., 1994). Our primary independent variable, *IPV perpetration*, was assessed via three items based on the Conflict Tactics Scale-2 (Straus, Hamby, Boney-McCoy, & Sugarman, 1996): physical IPV perpetration ("threatened partner with violence, pushed or shoved partner or threw something at partner"), sexual IPV perpetration ("insisted on or made partner have sex with him"), and causing injury due to IPV ("injured partner during a fight"). A subject was classified as "yes" for IPV perpetration if a positive response was recorded for any of the three items.

STI ever, our primary outcome, was assessed via survey items on whether the participant had ever been diagnosed with each of the following: syphilis, gonorrhea, Chlamydia, Trichomonas, genital warts, and genital herpes. *Current STI*, our secondary outcome, was assessed via biologic testing of urine and blood samples for gonorrhea, Chlamydia, Trichomonas, and syphilis (see Pace et al., in press for details).

In terms of sex risk behaviors, our secondary outcomes of interest, subjects provided detailed information about their sexual partners in the past 90 days (Kalichman et al., 2001). For each of their five most recent partners in this period, starting with the most recent, participants were asked about number of sexual episodes and condom use during these episodes, as well as number of episodes involving drug or alcohol use prior to sex (past 30 days). They were then asked about any other sex partners beyond five in the past 90 days using these same questions. These items were used to create the following additional secondary outcomes: *number of unprotected sex episodes* in the past 90 days, *multiple sex partners* (2+) in the past 90 days, and *substance use before sex* in the past 30 days. Sex trade involvement items were taken from the Risk Assessment Battery (RAB) (Navaline et al., 1994); one item asked about *buying sex* with money or drugs in the past 90 days.

Data analysis

Descriptive statistics were obtained for demographics, ART, sex risk behaviors, substance use behaviors, and STI for the overall sample and stratified by IPV perpetration status. Bivariate associations were assessed via chi-squared analyses, *t*-tests, or the nonparametric Wilcoxon rank sum test as appropriate. To assess associations between IPV perpetration and dichotomous STI and sex risk behavior outcomes, simple logistic regression analyses as well as multiple logistic regressions analyses adjusted for demographics (age, marital status, and education), ART, and substance use (binge alcohol use and IDU) were conducted. Crude and adjusted overdispersed Poisson regression analyses were used to assess associations between IPV perpetration and the number of unprotected sex episodes. Odds ratios (ORs) are reported for logistic regression models and incidence rate ratios (IRRs) are reported for Poisson regression models; 95% CIs were used to determine significance. Statistical analyses were performed by using SAS software (version 9.1; SAS Institute, Cary, NC, USA).

Results

Sample characteristics

Participants ($N = 415$) were aged 20–57 years; 20.7% had not completed high school, and 48.4% had a history of incarceration. Most of participants (81.0%) reported past 30-day binge alcohol use, and 42.7% had injected drugs in this same timeframe. Half of participants (50.1%) reported alcohol or drug use before sex in the past 30 days, and one in eight had bought sex in the past 90 days. More than 1 in 10 men (11.8%) currently had an STI; 41.0% had a history of a STI diagnosis. Almost half of participants (46%, $n = 189$) reported a history of IPV perpetration (Table 1). Of these IPV perpetrators, 55.9% ($n = 105/189$) reported physical IPV perpetration; 63.0% ($n = 119/189$) reported sexual IPV perpetration, and 46.9% ($n = 87/189$) reported injury due to IPV perpetration.

Associations between IPV and STI and sex risk behaviors

Adjusted regression analyses demonstrated that lifetime IPV perpetration is associated with a higher odds of reporting a lifetime STI (AOR = 1.58, 95% CI = 1.04, 2.39), but a lower odds of having a current STI (AOR = 0.50, 95% CI = 0.26, 0.96) (Table 2). No significant associations between lifetime IPV perpetration and sex risk behaviors were observed in adjusted analyses.

Discussion

Almost half of study participants reported a history of IPV perpetration, a percentage higher than that seen in other studies of men from Russia (Lysova & Hines, 2008; Serbanescu & Goodwin, 2005; WHO, 2011; Zhan et al., 2012), and may be attributable to having an HIV-infected and risky drinking sample. Previous research from Russia and elsewhere documents a strong association between IPV perpetration and substance use, particularly alcohol use (Lysova & Hines, 2008; WHO, 2011; Zhan et al., 2011, 2012). History of IPV perpetration may also be disproportionately represented among men with HIV in Russia, a finding seen in South Africa and India (Dunkle et al., 2006; Silverman et al., 2007). Of note, the proportion of IPV perpetration among HIV-infected substance users in Russia is consistent with findings from a sample of HIV-infected IDUs (Frye et al., 2007), highlighting the relevance of considerations of IPV perpetration in this population of men.

Table 1. Demographic and HIV/STI risk profile of male HIV-infected risky drinkers in St. Petersburg, Russia ($N = 415$) and stratified by for total sample ($N = 415$) and by history of IPV perpetration.

	Total sample $N = 415$	No history of IPV perpetration $N = 226$	History of IPV perpetration $N = 189$	p-Value*
Age range (years)	(20–57)	(20–54)	(22–57)	0.52
Mean (SD)	30.6 (4.8)	30.5 (4.7)	30.8 (4.8)	
Married	131 (31.6%)	63 (27.9%)	68 (36.2%)	0.07
Education				
≤10 grades	86 (20.7%)	55 (24.3%)	31 (16.4%)	0.05
11 + grades	329 (78.3%)	171 (75.7%)	158 (83.6%)	
Employed	345 (83.1%)	188 (83.2%)	157 (83.1%)	0.97
History of incarceration	201 (48.4%)	109 (48.2%)	92 (48.7%)	0.93
History of ART use	78 (19%)	43 (19%)	35 (19%)	0.90
Binge alcohol past 30 day	336 (81.0%)	183 (81.0%)	153 (81.0%)	0.99
IDU past 30 day	177 (42.7%)	106 (46.9%)	71 (37.6%)	0.06
MSM	15 (3.6%)	10 (4.4%)	5 (2.7%)	0.33
Multiple sex partners, past 90 days	127 (30.6%)	66 (29.2%)	61 (32.3%)	0.50
Buying sex	52 (12.5%)	25 (11.1%)	27 (14.3%)	0.32
Selling sex	12 (2.9%)	6 (2.7%)	6 (3.2%)	0.75
Any substance use before sex	208 (50.1%)	106 (46.9%)	102 (54.0%)	0.15
Number of unprotected sex episodes past 90 days range	(0–300)	(0–300)	(0–90)	0.002
Mean (SD)	13.5 (25.8)	12.7 (29.0)	14.4 (21.2)	
Median (IQR)	4.0 (1–15)	3 (0–10)	5 (2–17)	
Self-report STI ever	172 (41.0%)	82 (36.3%)	90 (47.6%)	0.02
Any current STI	48 (11.8%)	32 (14.6%)	16 (8.5%)	0.06

*p-Values based on chi-squared analyses, t-tests, or Wilcoxon test, as appropriate.

Table 2. Logistic and Poisson regression analyses assessing associations between IPV and the outcomes, STI and sex risk behaviors.

Outcomes	Crude OR (95% CI)	Adj[a] OR (95% CI)
Self-report STI ever	1.60 (1.08, 2.37) $p = 0.02$	1.58 (1.04, 2.39) $p = 0.03$
Current STI	0.55 (0.29, 1.03) $p = 0.06$	0.50 (0.26, 0.96) $p = 0.04$
Multiple sex partners	1.16 (0.76, 1.76) $p = 0.50$	1.24 (0.80, 1.93) $p = 0.33$
Buying sex	1.32 (0.75, 2.40) $p = 0.32$	1.36 (0.75, 2.48) $p = 0.31$
Alcohol or drug use before sex	1.33 (0.90, 1.96) $p = 0.15$	1.35 (0.89, 2.04) $p = 0.15$
	Crude IRR (95% CI)	Adj[a] IRR (95% CI)
Number of unprotected sex encounters	1.13 (0.78, 1.64) $p = 0.51$	1.06 (0.72, 1.58) $p = 0.76$

[a]Adjusted for age, marital status, education, ART, IDU, binge alcohol use.

Consistent with previous studies, findings indicate a higher odds of STI ever among those reporting a lifetime history of IPV perpetration. However, unexpectedly, history of IPV perpetration was not associated with current sexual risk behaviors and was negatively associated with current STI. These findings may in part be attributable to use of a lifetime IPV perpetration variable, which corresponds more directly to ever STI than current STI or recent sex risk behaviors, as such behaviors may change for men over time. Future research should examine whether recent IPV predicts current STI and sex risk behaviors in this population.

Study limitations include reliance on self-report of stigmatizing behaviors and STI over an "ever" timeframe; these data used are subject to recall and social desirability biases. Our IPV assessment had fewer items than the standard version, which may have resulted in underreporting. Use of differing time frames across some of our variables of interest (e.g., lifetime IPV perpetration, unprotected sex in past 90 days) likely affected some study findings. Cross-sectional data analyses do not allow an interpretation of the findings in terms of causality, and many potential relevant covariates such as knowledge and attitudes were not able to be considered by these analyses. Finally, the study may be limited in generalizability based on its focus on HIV-infected male risky drinkers recruited largely through health systems in the context of St. Petersburg, Russia. Despite limitations, study findings in conjunction with above described prior research document that IPV perpetration histories are a major concern among male HIV-infected risky drinkers in Russia and worthy of study with regard to their association with HIV acquisition and transmission in this national context.

Acknowledgements

The authors received grants from National Institutes of Health (R01AA016059, U24AA020778, U24AA020779, and K24AA015674).

References

Bogart, L.M., Collins, R.L., Cunningham, W., Beckman, R., Golinelli, D., Eisenman, D., & Bird, C.E. (2005). The association of partner abuse with risky sexual behaviors among women and men with HIV/AIDS. *AIDS and Behavior*, *9*(3), 325–333. doi:10.1007/s10461-005-9006-1

Decker, M.R., Miller, E., Kapur, N.A., Gupta, J., Raj, A., & Silverman, J.G. (2008). Intimate partner violence and sexually transmitted disease symptoms in a national sample of married Bangladeshi women. *International Journal of Gynaecology and Obstetrics*, *100*(1), 18–23. doi:10.1016/j.ijgo.2007.06.045

Decker, M.R., Seage, G.R., III, Hemenway, D., Gupta, J., Raj, A., & Silverman, J.G. (2009). Intimate partner violence perpetration, standard and gendered STI/HIV risk behaviour, and STI/HIV diagnosis among a clinic-based sample of men. *Sexually Transmitted Infections*, *85*(7), 555–560. doi:10.1136/sti.2009.036368

Dude, A. (2007). Intimate partner violence and increased lifetime risk of sexually transmitted infection among women in Ukraine. *Studies in Family Planning*, *38*(2), 89–100. doi:10.1111/j.1728-4465.2007.00120

Dunkle, K.L., Jewkes, R.K., Nduna, M., Levin, J., Jama, N., Khuzwayo, N., ... Duvvury, N. (2006). Perpetration of partner violence and HIV risk behaviour among young men in the rural Eastern Cape, South Africa. *AIDS*, *20*(16), 2107–2114. doi:10.1097/01.aids.0000247582.00826.52

El-Bassel, N., Fontdevila, J., Gilbert, L., Voisin, D., Richman, B.L., & Pitchell, P. (2001). HIV risks of men in methadone maintenance treatment programs who abuse their intimate partners: A forgotten issue. *Journal of Substance Abuse*, *13*(1–2), 29–43. doi:10.1016/S0899-3289(01)00068-2

El-Bassel, N., Gilbert, L., Wu, E., Go, H., & Hill, J. (2005a). HIV and intimate partner violence among methadone-maintained women in New York City. *Social Science & Medicine*, *61*(1), 171–183. doi:10.1016/j.socscimed.2004.11.035

El-Bassel, N., Gilbert, L., Wu, E., Go, H., & Hill, J. (2005b). Relationship between drug abuse and intimate partner

violence: A longitudinal study among women receiving methadone. *American Journal of Public Health*, *95*(3), 465–470. doi:10.2105/AJPH.2003.023200

El-Bassel, N., Witte, S.S., Wada, T., Gilbert, L., & Wallace, J. (2001). Correlates of partner violence among female street-based sex workers: Substance abuse, history of childhood abuse, and HIV risks. *AIDS Patient Care and STDs*, *15*(1), 41–51. doi:10.1089/108729101460092

Frye, V., Latka, M.H., Wu, Y., Valverde, E.E., Knowlton, A.R., Knight, K.R., … O'Leary, A. (2007). Intimate partner violence perpetration against main female partners among HIV-positive male injection drug users. *Journal of Acquired Immune Deficiency Syndromes*, *46*(Suppl. 2), S101–S109. doi:10.1097/QAI.0b013e31815767e6

Gilbert, L., El-Bassel, N., Wu, E., & Chang, M. (2007). Intimate partner violence and HIV risks: A longitudinal study of men on methadone. *Journal of Urban Health*, *84*(5), 667–680. doi:10.1007/s11524-007-9214-2

Goliusov, A.T., Dementyeva, L.A., Ladnaya, N.N., Briko, N.I., Tumanova, M.S., Korzhayeva, N.A., … Kobzeva, V.B. (2008). *Country progress report of the Russian Federation on the implementation of the declaration of commitment on HIV/AIDS*. Report period January 2006–December 2007. Retrieved from http://data.unaids.org/pub/Report/2008/Russia_2008_country_progress_report_en.pdf

Horne, S. (1999). Domestic violence in Russia. *The American Psychologist*, *54*(1), 55–61. doi:10.1037/0003-066X.54.1.55

Jewkes, R., Dunkle, K., Koss, M.P., Levin, J.B., Nduna, M., Jama, N., & Sikweyiya, Y. (2006). Rape perpetration by young, rural South African men: Prevalence, patterns and risk factors. *Social Science & Medicine*, *63*(11), 2949–2961. doi:10.1016/j.socscimed.2006.07.027

Jewkes, R., Dunkle, K., Nduna, M., Levin, J., Jama, N., Khuzwayo, N., … Duvvury, N. (2006). Factors associated with HIV sero-status in young rural South African women: Connections between intimate partner violence and HIV. *International Journal of Epidemiology*, *35*(6), 1461–1468. doi:10.1093/ije/dyl218

Kalichman, S.C., Kelly, J.A., Shaboltas, A., & Granskaya, J. (2000). Violence against women and the impending AIDS crisis in Russia. *The American Psychologist*, *55*(2), 279–280. doi:10.1037/0003-066X.55.2.279

Kalichman, S.C., Rompa, D., Cage, M., DiFonzo, K., Simpson, D., Austin, J., … Graham, J. (2001). Effectiveness of an intervention to reduce HIV transmission risks in HIV-positive people. *American Journal of Preventive Medicine*, *21*(2), 84–92. doi:10.1016/S0749-3797(01)00324-5

Lokhmatkina, N.V., Kuznetsova, O.Y., & Feder, G.S. (2010). Prevalence and associations of partner abuse in women attending Russian general practice. *Family Practice*, *27*(6), 625–631. doi:10.1093/fampra/cmq044

Lysova, A.V., & Hines, D.A. (2008). Binge drinking and violence against intimate partners in Russia. *Aggressive behavior*, *34*(4), 416–427. doi:10.1002/ab.20256

Navaline, H.A., Snider, E.C., Petro, C.J., Tobin, D., Metzger, D., Alterman, A.I., & Woody, G.E. (1994). Preparations for AIDS vaccine trials. An automated version of the Risk Assessment Battery (RAB): Enhancing the assessment of risk behaviors. *AIDS Research and Human Retroviruses*, *10*(Suppl. 2), S281–S283.

Pace, C., Lioznov, D., Cheng, D.M., Wakeman, S.M., Raj, A., Walley, A.Y., … Samet, J.H. (in press). Sexually transmitted infections among HIV-infected heavy drinkers in St. Petersburg, Russia. *International Journal of STD & AIDS*.

Panchanadeswaran, S., Frye, V., Nandi, V., Galea, S., Vlahov, D., & Ompad, D. (2010). Intimate partner violence and consistent condom use among drug-using heterosexual women in New York City. *Women Health*, *50*(2), 107–124. doi:10.1080/03630241003705151

Petry, N.M. (2001). Reliability of drug users' self-reported HIV risk behaviors using a brief, 11-item scale. *Substance Use & Misuse*, *36*(12), 1731–1747. doi:10.1081/JA-100107576

Raj, A., Reed, E., Welles, S.L., Santana, M.C., & Silverman, J.G. (2008). Intimate partner violence perpetration, risky sexual behavior, and STI/HIV diagnosis among heterosexual African American men. *American Journal of Men's Health*, *2*(3), 291–295. doi:10.1177/1557988308320269

Raj, A., Santana, M.C., La Marche, A., Amaro, H., Cranston, K., & Silverman, J.G. (2006). Perpetration of intimate partner violence associated with sexual risk behaviors among young adult men. *American Journal of Public Health*, *96*(10), 1873–1878. doi:10.2105/AJPH.2005.081554

Santana, M.C., Raj, A., Decker, M.R., La Marche, A., & Silverman, J.G. (2006). Masculine gender roles associated with increased sexual risk and intimate partner violence perpetration among young adult men. *Journal of Urban Health*, *83*(4), 575–585. doi:10.1007/s11524-006-9061-6

Schmuel, E., & Schenker, J.G. (1998). Violence against women: The physician's role. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*, *80*(2), 239–245. doi:10.1016/S0301-2115(98)00140-7

Serbanescu, F., & Goodwin, M. (2005, September). *Domestic violence in Eastern Europe: Levels, risk factors and selected reproductive health consequences*. Paper presented at the European Conference on Interpersonal Violence, Paris, France.

Silverman, J.G., Decker, M.R., Kapur, N.A., Gupta, J., & Raj, A. (2007). Violence against wives, sexual risk and sexually transmitted infection among Bangladeshi men. *Sexually Transmitted Infections*, *83*(3), 211–215. doi:10.1136/sti.2006.023366

Silverman, J.G., McCauley, H.L., Decker, M.R., Miller, E., Reed, E., & Raj, A. (2011). Coercive forms of sexual risk and associated violence perpetrated by male partners of female adolescents. *Perspectives on Sexual and Reproductive Health*, *43*(1), 60–65. doi:10.1363/4306011

Silverman, J.G., Raj, A., Mucci, L.A., & Hathaway, J.E. (2001). Dating violence against adolescent girls and

associated substance use, unhealthy weight control, sexual risk behavior, pregnancy, and suicidality. *JAMA*, *286*(5), 572–579. doi:10.1001/jama.286.5.572

Sobell, L.C., & Sobell, M.B. (1992). Timeline follow-back: A technique for assessing self-reported alcohol consumption. In R.Z. Litten & J.P. Allen (Eds.), *Measuring alcohol consumption: Psychosocial and biological methods* (pp. 41–72). Totowa, NJ: Humana Press.

Sobell, L.C., & Sobell, M.B. (1996). *Alcohol timeline followback (TFLB) users' manual*. Toronto: Addiction Research Foundation.

Straus, M.A., Hamby, S.L., Boney-McCoy, S., & Sugarman, D.B. (1996). The revised conflict tactics scales (CTS2). *Journal of Family Issues*, *17*(2), 283–316. doi:10.1177/019251396017003001

Weatherby, N.L., Needle, R., Cesari, H., Booth, R., McCoy, C.B., Watters, J.K., ... Chitwood, D.D. (1994). Validity of self-reported drug use among injection drug users and crack cocaine users recruited through street outreach. *Evaluation and Program Planning*, *17*(4), 347–355. doi:10.1016/0149-7189(94)90035-3

World Health Organization (WHO). (2005). *Summary country profile for HIV/AIDS treatment scale-up: Russian Federation*. Retrieved from http://www.who.int/hiv/HIVCP_RUS.pdf

World Health Organization (WHO). (2011). *Intimate partner violence and alcohol fact sheet*. Retrieved from http://www.who.int/violence_injury_prevention/violence/world_report/factsheets/ft_intimate.pdf

Zhan, W., Hansen, N.B., Shaboltas, A.V., Skochilov, R.V., Kozlov, A.P., Krasnoselskikh, T.V., & Abdala, N. (2012). Partner violence perpetration and victimization and HIV risk behaviors in St. Petersburg, Russia. *Journal of Traumatic Stress*, *25*(1), 86–93. doi:10.1002/jts.21658

Zhan, W., Shaboltas, A.V., Skochilov, R.V., Kozlov, A.P., Krasnoselskikh, T.V., & Abdala, N. (2011). Alcohol misuse, drinking contexts and intimate partner violence in St. Petersburg, Russia: Results from a cross-sectional study. *BMC Public Health*, *11*, 629. doi:10.1186/1471-2458-11-629

The Oncologist®

Breast Cancer

Patient Navigation for Underserved Patients Diagnosed with Breast Cancer

APARNA RAJ,[a] NAOMI KO,[b] TRACY A. BATTAGLIA,[b] BRUCE A. CHABNER,[a] BEVERLY MOY[a]

[a]Massachusetts General Hospital, Boston, Massachusetts, USA; [b]Boston Medical Center, Boston, Massachusetts, USA

Key Words. Patient navigation • Cancer disparities • Breast cancer • Outcomes

Disclosures: Bruce A. Chabner: BOD - PharmaMar (Spain) BOD - CTNET SAB- Epizyme (E); Merrimack, Sanofi, Peregrine, Allergan, Redwood (C/A); **Beverly Moy:** Avon Foundation (RF). The other authors indicated no financial relationships.

Section Editors: **Gabriel Hortobágyi**: Amgen, Antigen Express, Galena Biopharma, Genentech, Novartis, Rockpointe (C/A); Novartis (RF); Taivex (O); Citizen's Oncology Foundation (founder and member of the board of directors); **Kathleen Pritchard**: Novartis, Roche, AstraZeneca, Pfizer, Abraxis, Boehringer-Ingelheim, GlaxoSmithKline (C/A, RF)

Reviewer "A": Expert Advisory Panel for Sanofi, Gerson Lehrman (C/A)

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

LEARNING OBJECTIVES

After completing this course, the reader will be able to:

1. Describe the role and potential benefits of patient navigation in breast cancer care.
2. Explain disparities in breast cancer care and their impact on patient populations.

CME This article is available for continuing medical education credit at CME.TheOncologist.com.

ABSTRACT

The elimination of cancer disparities is critically important for lessening the burden of breast cancer (BC). Patient navigator programs (PNPs) have been shown to improve rates of BC screening in underserved communities, but there is a dearth of evidence regarding their benefits after the actual diagnosis of BC. We retrospectively examined sociodemographic characteristics, disease characteristics, and concordance to quality measures (QMs) of BC care among women participating in a PNP that services disadvantaged minority communities in the greater Boston area. Of the 186 PNP patients diagnosed with BC in 2001–2011 in three neighborhood community health centers, treatment data was available for 158 (85%) and race and disease stage information was available for 149 (80%). Regarding stage, 25% were diagnosed with in situ cancer, 32% had stage 1, 25% had stage 2, 13% had stage 3, and 5% had stage 4 BC. Guideline-indicated care was received by 70 of 74 patients (95%) for the hormonal therapy QM, 15 of 17 (88%) patients for the chemotherapy QM, and 65 of 71 (92%) patients for the radiation QM, all similar to published concordance rates at elite National Comprehensive Cancer Network institutions. These findings suggest that PNPs may facilitate evidence-based quality care for vulnerable populations. Future research should prospectively analyze quality metrics to assess measures to improve the process and outcomes of patient navigation in diverse underserved settings, compared with control non-navigated populations. ***The Oncologist*** 2012;17:1027–1031

Correspondence: Beverly Moy, M.D., M.P.H., Gillette Center for Women's Cancers, Harvard Medical School, Massachusetts General Hospital, 55 Fruit Street, Yawkey 9A, Boston, Massachusetts 02114, USA. Telephone: 617-724-1074; Fax: 617-724-1079; e-mail: bmoy@partners.org Received May 2, 2012; accepted for publication June 12, 2012; first published online in ***The Oncologist*** *Express* on July 2, 2012. 1083-7159/2012/$20.00/0 http://dx.doi.org/10.1634/theoncologist.2012-0191

Background

Despite major advances in cancer research, screening, and treatment, not all Americans with cancer have benefited equally. Although there was a 14% decrease in cancer-related deaths over the years 1991–2004, racial and ethnic minority patients continue to die disproportionately from cancer, compared with their white counterparts, even after adjusting for insurance status and income [1]. There is increasing evidence that the disconnect between discoveries in cancer care and their timely delivery to all Americans contributes to cancer disparities. Solutions to improve the equity of cancer care delivery are desperately needed [2].

Patient navigation programs (PNPs) have emerged as a potential solution for improving cancer care delivery [3, 4]. PNPs facilitate access to quality medical care by identifying barriers to care and by bridging gaps in care through culturally sensitive coordination. Patient navigators are resources for patients and providers and may assist with all phases of access, including primary cancer prevention, screening and follow-up care, cancer treatment, and survivorship care [5].

Extensive data have established the efficacy of navigation in improving the timeliness and receipt of cancer screening and diagnostic care after an abnormal screening test [6–11]. However, it is unknown whether or not PNPs improve patient care and outcomes following the actual diagnosis of cancer. We sought to evaluate the clinical outcomes of patients enrolled in the Massachusetts General Hospital (MGH) Avon Breast Care Program (MABCP) based on evidence-based national quality measures. The MABCP provides patient navigation services to racially and ethnically diverse communities seeking care at four federally qualified health centers in the greater Boston area, and thus represent a population vulnerable for poor cancer outcomes.

Methods

We performed a retrospective chart review of all 186 women diagnosed with breast cancer who participated in the MABCP in 2001–2011. Since its inception in 2001, the MABCP has served four community health centers and has provided primary screening services and diagnostic follow-up of abnormal screening examinations for ~4,000 patients. Of these 4,000 patients, 186 patients (4.7%) were diagnosed with breast cancer.

Patients in need of routine screening mammography are referred to the MABCP by their community health center primary care physician. The patient navigators automatically enroll all patients at these centers who have abnormal screening examinations and require diagnostic follow-up examinations. MABCP navigators follow the community health worker model [5] and are trained lay workers who are culturally diverse and generally representative of the population served by their community health center. Languages spoken by the navigators include Spanish, Bosnian, and Portuguese.

Most patients received their subsequent cancer care at either MGH or Boston Medical Center. Institutional review board approval was obtained at both sites. Sociodemographic data and treatment data were obtained by review of patient charts within the respective electronic medical records.

Table 1. Sociodemographic characteristics of Massachusetts General Hospital Avon Breast Care Program breast cancer patients ($n = 186$)

Characteristic	*n*	%
Race or ethnicity		
White	59	32
Black	30	16
Hispanic	52	28
Asian/Middle Eastern	8	4
Not disclosed	37	20
Language		
English	106	57
Spanish	55	30
Creole	10	5
Vietnamese	10	5
Bosnian	3	2
Portuguese	2	1
Median age (range), yrs	58 (19–93)	
Insurance status		
Uninsured	12	6
Free Care	5	3
Medicaid	23	12
Medicare	45	24
Private	58	31
Unknown	43	23
Highest education level		
None	2	1
≤Grade 8	25	13
Some high school	12	6
High school or general educational development	48	26
Some college	9	5
College graduate	12	6
Unknown	82	44

Descriptive statistics were used to analyze baseline sociodemographic characteristics. Clinical outcomes were analyzed using data abstracted from electronic medical records. Two physicians collected data using a chart abstraction form and entered the data into a secure Microsoft® Excel database. Abstracted clinical data were used to examine concordance with American Society of Clinical Oncology/National Comprehensive Cancer Network (ASCO/NCCN) quality measures.

Quality Measures

Three ASCO/NCCN quality measures based on NCCN level 1 evidence were analyzed to determine quality of cancer care. These included: (a) hormonal therapy within 1 year of diagnosis of hormone receptor $(HR)^+$ tumors >1 cm, (b) chemother-

Table 2. Race or ethnicity and disease stage of Massachusetts General Hospital Avon Breast Care Program breast cancer patients ($n = 149$)

Race or ethnicity	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4	Race, *n* (%)
White	16	18	17	5	3	59 (40%)
Black	5	7	8	9	1	30 (20%)
Hispanic	16	19	11	3	3	52 (35%)
Asian or Middle Eastern	0	4	2	2	0	8 (5%)
Total (%)	37 (25%)	48 (32%)	38 (25%)	19 (13%)	7 (5%)	149

Table 3. Concordance rates with American Society of Clinical Oncology/NCCN guidelines (MABCP versus NCCN centers)

Quality measure	MABCP	NCCN Centers	*p*-value
Hormonal therapy	95% (70/74)	89% (2751/3091)	.72
Chemotherapy	88% (15/17)	87% (1044/1200)	.97
Radiation therapy	92% (65/71)	95% (2679/2829)	.85

Abbreviations: MABCP, Massachusetts General Hospital Avon Breast Care Program; NCCN, National Comprehensive Cancer Network.

apy within 120 days of diagnosis of HR^- tumors >1 cm for women aged <70 years, and (c) postlumpectomy radiation therapy.

Descriptive statistics were used to report patient characteristics and determine the proportion of navigated patients who had documented care that was in concordance with these three quality metrics. Concordance data for these quality measures across all NCCN institutions were used as a benchmark for MABCP patients [12]. Concordance rates between MABCP patients and NCCN institution patients were compared using χ^2 analyses.

Results

Self-reported sociodemographic data, including language spoken, age, insurance status, and level of education of MABCP patients diagnosed with breast cancer was available for the 186 patients in the MABCP. Missing data were attributed to patients declining to report these characteristics. In terms of race and ethnicity, 60% of the patients in the program were nonwhite black, Hispanic, and Asian or Middle Eastern. Forty-three percent of MABCP patients were non–English speaking. Twenty-one percent were either uninsured or were enrolled in Medicaid or Free Care. Thirty-seven percent had high school or general educational development education or greater (Table 1).

Of the 186 MABCP patients diagnosed with breast cancer, treatment data were available for 158 (85%) and race and breast cancer stage information was available for 149 (80%) (Table 2). In terms of stage of disease at diagnosis, 25% were diagnosed with in situ cancer, 32% had stage 1, 25% had stage 2, 13% had stage 3, and 5% had stage 4 breast cancer. Although there were no obvious differences between racial or ethnic groups with regard to breast cancer stage, the numbers are too small to make direct statistical comparisons. In our limited dataset, black patients appeared to be diagnosed with stage 3 disease more frequently than white patients. In terms of disease characteristics (data not shown), 91% of the navigated patients had HR^+ disease and 10% of the patients had human epidermal growth factor receptor 2–positive disease.

QMs

Table 3 illustrates concordance with ASCO/NCCN quality measure among MABCP patients compared with concordance rates for patients who received breast cancer care at NCCN institutions. For the hormonal therapy quality measure, 95% of the MABCP patients received hormonal therapy within 1 year of diagnosis for HR^+ tumors >1 cm. For the chemotherapy quality measure, 88% of the patients in the MABCP received chemotherapy within 120 days of diagnosis of HR^- breast cancer >1 cm. For the radiation quality measure, 92% of the patients received postlumpectomy radiation therapy. These concordance rates are comparable with rates from eight NCCN centers, largely viewed as elite cancer centers in the U.S. [12]. There was no significant difference ($p > .05$) between MABCP patients and NCCN patients with regard to each of the quality measures. Upon detailed review of the treatment records of MABCP patients, patients received standard adjuvant chemotherapy regimens such as doxorubicin, cyclophosphamide followed by paclitaxel, and docetaxel plus cyclophosphamide. With regard to endocrine therapy, the proper medication was prescribed to patients. For example, tamoxifen was prescribed for premenopausal patients and aromatase inhibitors were prescribed for postmenopausal patients.

Discussion

In this study, we demonstrated that breast cancer patients who received navigation services received high-quality cancer care, as defined by concordance with ASCO/NCCN quality measures. These navigated patients also had a favorable breast cancer stage distribution, with >50% having in situ or stage 1 disease, similar to that of white women reported by the Surveillance, Epidemiology, and End Results program [13]. This staging profile is also comparable with that of the white breast cancer population in Massachusetts as reported by the Massachusetts Cancer Registry [14, 15]. Our findings are particularly

notable given the sociodemographic characteristics of these patients, a significant proportion of whom are non–English speaking, are underinsured, and have a low level of education. Our findings add to the growing body of literature that suggests that navigation is a viable approach to reduce breast cancer health disparities by improving early detection rates and perhaps ensuring receipt of quality cancer treatment.

Much of the patient navigation literature addressing the care of cancer patients focuses on improving screening rates and the diagnostic management of abnormal screening results [9, 16, 17]. Most studies have shown better outcomes when patients receive patient navigation services. However, evidence that PNPs improve clinical outcomes after a cancer diagnosis is lacking. Arguably, patient navigation is equally, if not more, important after a cancer diagnosis given the logistical complexities and financial burdens of cancer treatment such as chemotherapy or radiation. In our patient population, 15% of the patients opted for treatment at other institutions, and therefore their treatment data were unavailable. However, among those patients with available data, at least 87% received care that adhered to guidelines. Some reports suggest that patient navigation after a cancer diagnosis improves patient satisfaction and reduces barriers to care [18, 19]. Guadagnolo and colleagues showed that patient navigation led to fewer cancer treatment interruptions and higher rates of clinical trial enrollment among Native American cancer patients, compared with historical controls [20]. Ell and colleagues showed that cancer treatment adherence was better than in historical controls among patients who received navigation services [21]. Our study offers additional data to support the efficacy of PNPs for patients diagnosed with cancer.

Our finding of a high concordance with level 1 ASCO/NCCN quality measures has important implications for the underserved and vulnerable patients diagnosed with breast cancer. These level 1 quality measures are based on randomized controlled clinical trials. Our data suggest that patient navigation may offer an effective strategy to overcome barriers that interfere with access to quality cancer care for underserved communities.

Major weaknesses of our study are that it was a retrospective analysis and data regarding the outcomes of 15% of the patients are missing. Fifteen percent of MABCP patients elected to receive their cancer care at institutions other than MGH or Boston Medical Center, and therefore we lack treatment data for these patients. However, their sociodemographic data otherwise closely matched those of the patients for whom we do have treatment data, suggesting that their treatment experience may match that of the larger population. Another weakness of our study is the extent of missing data, particularly with respect to education and insurance status. Our cohort may be disproportionately influenced by more highly educated and insured patients, leading to selection bias. Another weakness of this study is the lack of a closely matched control group that did not receive patient navigation. Similar studies in the literature often use historical controls as a reference group, but this strategy also has weaknesses, including the inability to account for improved outcomes with modern cancer therapies. We used data from patients receiving treatment at NCCN centers during the same time period as the benchmark to demonstrate that our navigated patients received similar quality care. Finally, another weakness of the study is the relatively small number of women diagnosed with breast cancer ($n = 186$) within the MABCP, making it difficult to draw definitive conclusions about this population.

Despite these weaknesses, our findings provide objective evidence that vulnerable populations may indeed receive quality cancer care when enrolled in a PNP, further prioritizing the need for rigorous research, ideally large, multicenter, prospective, randomized, controlled trials of patient navigation in patients diagnosed with cancer. Such trials are currently under way at the nine sites participating in the National Cancer Institute (NCI) Patient Navigation Research Program [22]. However, these trials are primarily enrolling patients for cancer screening, not after a cancer diagnosis. Additionally, the issue of randomizing patients to an intervention that is intuitively beneficial versus control (no intervention) raises ethical considerations. However, this national effort will provide insight into the role of patient navigation after a breast cancer diagnosis and, importantly, will allow evaluation of the cost-effectiveness of this approach [23, 24].

In conclusion, we have shown that patient navigation can lead to high-quality breast cancer care, as measured by concordance with national guidelines, among an underserved vulnerable population in an urban setting. Further research is needed to examine other metrics of quality care that may be improved with patient navigation during cancer treatment.

Author Contributions

Conception/Design: Beverly Moy, Aparna Raj, Naomi Ko, Tracy A. Battaglia
Provision of study material or patients: Beverly Moy
Collection and/or assembly of data: Beverly Moy, Aparna Raj, Naomi Ko
Data analysis and interpretation: Beverly Moy, Aparna Raj, Naomi Ko, Tracy A. Battaglia, Bruce A. Chabner
Manuscript writing: Beverly Moy, Aparna Raj, Naomi Ko, Tracy A. Battaglia, Bruce A. Chabner
Final approval of manuscript: Beverly Moy, Aparna Raj, Naomi Ko, Tracy A. Battaglia, Bruce A. Chabner

References

1. Mead H, Cartwright-Smith L, Jones K et al. Racial and Ethnic Disparities in US Healthcare: A Chartbook. New York: The Commonwealth Fund, 2008.

2. Moy B, Polite BN, Halpern MT et al. American Society of Clinical Oncology policy statement: Opportunities in the patient protection and affordable care act to reduce cancer care disparities. J Clin Oncol 2011;29:3816–3824.

3. Guadagnolo BA, Dohan D, Raich P. Metrics for evaluating patient navigation during cancer diagnosis and treatment: Crafting a policy-relevant research agenda for patient navigation in cancer care. Cancer 2011;117(15 suppl):3565–3574.

4. Moy B, Chabner BA. Patient navigator programs, cancer disparities, and the patient protection and affordable care act. ***The Oncologist*** 2011;16:926–929.

5. Freeman HP, Rodriguez RL. History and principles of patient navigation. Cancer 2011;117(15 suppl):3539–3542.

6. Fowler T, Steakley C, Garcia AR et al. Reducing disparities in the burden of cancer: The role of patient navigators. PLoS Med 2006;3:e193.

7. Freeman HP. A model navigation program. Oncol Iss 2004;Sept/Oct:44–46.

8. Schwaderer KA, Proctor JW, Martz EF et al. Evaluation of patient navigation in a community radiation on-

cology center involved in disparities studies: A time-to-completion-of-treatment study. J Oncol Pract 2008;4: 220–224.

9. Wells KJ, Battaglia TA, Dudley DJ et al. Patient navigation: State of the art or is it science? Cancer 2008; 113:1999–2010.

10. Battaglia TA, Roloff K, Posner MA et al. Improving follow-up to abnormal breast cancer screening in an urban population. A patient navigation intervention. Cancer 2007;109(2 suppl):359–367.

11. Phillips CE, Rothstein JD, Beaver K et al. Patient navigation to increase mammography screening among inner city women. J Gen Intern Med 2011;26:123–129.

12. Hughes ME, Ottesen R, Niland JC et al. Quality of breast cancer care in NCCN centers as assessed by the ASCO/NCCN quality measures: Overall performance and reasons for nonconcordance [abstract 6506]. J Clin Oncol 2009;27(15 suppl):324S.

13. National Cancer Institute. Surveillance, Epidemiology, and End Results. SEER Cancer Statistics Review, 1973–1998. Available at http://seer.cancer.gov/csr/1973_1998/, accessed April 24, 2012.

14. Lobb R, Ayanian JZ, Allen JD et al. Stage of breast cancer at diagnosis among low-income women with access to mammography. Cancer 2010;116:5487–5496.

15. Massachusetts Cancer Registry. Cancer in Massachusetts by Race and Ethnicity, 2000–2004. Available at http://www.mass.gov/eohhs/docs/dph/cancer/race-ethnicity.pdf, accessed April 10, 2012.

16. Christie J, Itzkowitz S, Lihau-Nkanza I et al. A randomized controlled trial using patient navigation to increase colonoscopy screening among low-income minorities. J Natl Med Assoc 2008;100:278–284.

17. Dignan MB, Burhansstipanov L, Hariton J et al. A comparison of two Native American Navigator formats: Face-to-face and telephone. Cancer Control 2005; 12(suppl 2):28–33.

18. Campbell C, Craig J, Eggert J et al. Implementing and measuring the impact of patient navigation at a comprehensive community cancer center. Oncol Nurs Forum 2010;10:61–68.

19. Carroll JK, Humiston SG, Meldrum SC et al. Patients' experiences with navigation for cancer care. Patient Educ Couns 2010;80:241–247.

20. Guadagnolo BA, Boylan A, Sargent M et al. Patient navigation for American Indians undergoing cancer treatment: Utilization and impact on care delivery in a regional healthcare center. Cancer 2011;117:2754–2761.

21. Ell K, Vourlekis B, Xie B et al. Cancer treatment adherence among low-income women with breast or gynecologic cancer: A randomized controlled trial of patient navigation. Cancer 2009;115:4606–4615.

22. National Cancer Institute. Center to Reduce Cancer Health Disparities. Patient Navigation Research Program (PNRP). Available at http://crchd.cancer.gov/pnp/pnrp-index.html, accessed March 28, 2012.

23. Freund KM, Battaglia TA, Calhoun E et al. National Cancer Institute Patient Navigation Research Program: Methods, protocol, and measures. Cancer 2008; 113:3391–3399.

24. Ramsey S, Whitley E, Mears VW et al. Evaluating the cost-effectiveness of cancer patient navigation programs: Conceptual and practical issues. Cancer 2009; 115:5394–5403.

Published in final edited form as:
Clin Trials. 2012 April ; 9(2): 176–187. doi:10.1177/1740774511433284.

Analysis of Combined Data from Heterogeneous Study Designs: A Methodological Proposal from the Patient Navigation Research program

Richard G. Roetzheim, MD, MSPH[a], **Karen M. Freund, MD, MPH**[b], **Don K. Corle, MS**[c], **David M. Murray, PHD**[d], **Frederick R. Snyder, ABD**[e], **Andrea C. Kronman, MD, MSc**[b], **Pascal Jean-Pierre, PhD, MPH**[f], **Peter C. Raich, MD**[g], **Alan E. C. Holden, PhD**[h], **Julie S. Darnell, PhD, MHSA**[i], **Victoria Warren-Mears, PhD, RD**[j], **Steven Patierno, PhD**[k], **PNRP Design**, and **Analysis Committee For the Patient Navigation Research Program Investigators**

[a]U. of South Florida Department of Family Medicine and the H. Lee Moffitt Cancer Center, Tampa, FL, USA

[b]Women's Health Unit, Department of Medicine, Boston University School of Medicine, Boston, MA, USA

[c]National Cancer Institute, Rockville, MD, USA

[d]Division of Epidemiology, College of Public Health, The Ohio State University, Columbus, OH, USA

[e]Nova Research Company, Bethesda, MD, USA

[f]University of Miami Miller School of Medicine and Sylvester Comprehensive Cancer Center, Miami, FL, USA

[g]Division of Hematology/Oncology, Denver Health and Hospital Authority, Denver, CO, USA

[h]Institute for Health Promotion Research, The University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

[i]Division of Health Policy and Administration, School of Public Health, University of Illinois at Chicago, Chicago, IL, USA

[j]Northwest Tribal Epidemiology Center, Northwest Portland Area Indian Health Board, Portland, OR, USA

[k]George Washington Cancer Institute, George Washington University Medical Center, Washington, DC, USA

Abstract

Background—The Patient Navigation Research Program (PNRP) is a cooperative effort of nine research projects, each employing its own unique study design. To evaluate projects such as PNRP, it is desirable to perform a pooled analysis to increase power relative to the individual projects. There is no agreed upon prospective methodology, however, for analyzing combined data arising from different study designs. Expert opinions were thus solicited from members of the PNRP Design and Analysis Committee

Address for Correspondence: Richard Roetzheim, MD, MSPH, University of South Florida, Department of Family Medicine, 12901 Bruce B. Downs Blvd., MDC 13, Tampa, FL 33612, (813) 974-2445, (813) 974-4057 fax, rroetzhe@health.usf.edu.

[l]Other members of the PNRP Design and Analysis Committee are Joe Feinglass, Beth Calhoun, Diane Fairclough, Betsy Risendal, Victoria Warren-Mears, Kevin Fiscella, Ji-Hyun Lee, Nancy LaVerda, Paul Levine, Heather Hoffman, Ken Chu, Martha Hare, Mary Ann Van Duyn, Mollie Howerton

Purpose—To review possible methodologies for analyzing combined data arising from heterogeneous study designs.

Methods—The Design and Analysis Committee critically reviewed the pros and cons of five potential methods for analyzing combined PNRP project data. Conclusions were based on simple consensus. The five approaches reviewed included: 1) Analyzing and reporting each project separately, 2) Combining data from all projects and performing an individual-level analysis, 3) Pooling data from projects having similar study designs, 4) Analyzing pooled data using a prospective meta analytic technique, 5) Analyzing pooled data utilizing a novel simulated group randomized design.

Results—Methodologies varied in their ability to incorporate data from all PNRP projects, to appropriately account for differing study designs, and in their impact from differing project sample sizes.

Limitations—The conclusions reached were based on expert opinion and not derived from actual analyses performed.

Conclusions—The ability to analyze pooled data arising from differing study designs may provide pertinent information to inform programmatic, budgetary, and policy perspectives. Multi-site community-based research may not lend itself well to the more stringent explanatory and pragmatic standards of a randomized controlled trial design. Given our growing interest in community-based population research, the challenges inherent in the analysis of heterogeneous study design are likely to become more salient. Discussion of the analytic issues faced by the PNRP and the methodological approaches we considered may be of value to other prospective community-based research programs.

Keywords

Patient navigation; Health disparities; Pooled analysis; Research methodology

Introduction

Patient navigation is a promising approach to reduce cancer disparities and refers to support and guidance offered to persons with abnormal cancer screening or a new cancer diagnosis in order to more effectively access the cancer care system. [1] The primary goals of navigation are to help patients overcome barriers to care and facilitate timely, quality care provided in a culturally sensitive manner. Patient navigation is intended to target those who are most at risk for delays in care, including individuals from racial and ethnic minority and lower income populations.

Although patient navigation is an intervention that has clearly grown in popularity over the past decade, rigorous research on the efficacy of patient navigation is still new.[2] A number of studies conducted on patient navigation have been recently summarized.[2, 3] Patient navigation has generally shown improvements in timeliness of definitive diagnosis and initiation of cancer care. To date, there have been no published multicenter studies assessing patient navigation.

The Patient Navigation Research Program (PNRP) is the first multi-center program to critically examine the role and benefits of patient navigation. This program is sponsored and funded by the National Cancer Institute's (NCI) Center to Reduce Cancer Health Disparities (CRHCD), with additional support from the American Cancer Society. The five-year program focuses on four common cancers (breast, cervical, colorectal, and prostate) with screening tests having evidence of disparate outcomes in underserved populations.

The PNRP is a cooperative effort of nine research projects (see Figure 1), the funding agencies, and an evaluation contractor. Each PNRP project focuses on one or more of the four specified cancers. All projects targeted populations at greater risk of disparate cancer outcomes, such as racial or ethnic minorities, the uninsured or underinsured, or persons of lower socioeconomic status. As a cooperative endeavor, the PNRP utilizes a steering committee consisting of the nine project principal investigators, along with representatives from NCI-CRCHD, the American Cancer Society, and NOVA Research Company (NOVA), the evaluation contractor for the national outcomes studies. In order to advise the steering committee on methodological issues, a design and analysis committee was created composed of investigators having expertise in research design and methods from each project site, NOVA, and NCI-CRCHD.

The cooperative agreement did not require a uniform research design across all projects. This allowed for flexibility in implementing the patient navigation intervention that would be sensitive both to the specific patient populations as well as local system-level factors at each project. In addition, the nature of patient navigation requires involvement of the community with development of the research strategy in collaboration with community partners. Thus, each project had its own research design that included traditional randomized clinical trials (RCT), group-randomized trials (GRT), and non-randomized quasi-experimental designs (QE).

Despite the differences in research designs, all projects included characteristics specified by the steering committee, in order to support a single common evaluation of the PNRP. For example, all projects utilized a common definition of patient navigation and navigators received common national training in this role. In addition, all study designs share common well-defined outcomes with data elements from a single detailed data dictionary.

Navigation was hypothesized to shorten the time interval from cancer screening abnormality to definitive diagnosis (primary outcome) and for patients diagnosed with cancer, the time interval from definitive diagnosis through initial cancer treatment as well as improving satisfaction with care. Each project assessed these time intervals by medical record abstraction and the common data elements from all projects were uploaded to a national database to support a single common evaluation of the PNRP.

While each PNRP project was individually powered to address its hypotheses, a pooled analysis was desirable for many reasons. First, the increased statistical power from pooling data allows a more precise estimate of navigation effects across a variety of settings. In addition, while cancer screening abnormalities are common in primary medical care settings, cancer diagnoses are rare. The ability to address the hypothesis that navigation improves cancer care is strengthened by pooling data across projects to increase statistical power. Pooling of data also increases statistical power to explore navigation effects within patient subgroups at greater risk of adverse cancer outcomes (uninsured, racial-ethnic minorities, and persons of lower socioeconomic status). Pooling increases the generalizability of the findings by demonstrating the effect across a variety of settings and populations. Finally, pooling data enables exploration of heterogeneity of navigation effects (contrasting models of navigation for example), something that is not possible within individual PNRP projects.

For these reasons, a pooled analysis to evaluate the overall PNRP was desirable. The design and structure of PNRP, however, created a challenge in developing a suitable analytic method to evaluate the overall program. The PNRP was clearly different from standard multicenter trials that utilize a single study design with common shared protocols at all centers. In addition, while there was a common definition of navigation, its delivery could

differ among the PNRP projects. The design variations across the distinct projects do not lend themselves to traditional methods for analyzing data from a multi-center research trial.

Other fields, notably epidemiology, have recognized the need to analyze pooled data arising from heterogeneous study designs.[4, 5] These efforts generally involve retrospective analysis of published studies. To our knowledge, however, there is no agreed upon methodology for analyzing combined data arising from different study designs a priori or before each project has published their primary results, as opposed to retrospective analyses. The Design and Analysis Committee was therefore charged with the task of reviewing potential methodologies for analyzing combined data from PNRP projects. This review took place over a series of conference calls and the results of this assessment are presented below. In the following section, we describe each of the PNRP projects in more detail, and discuss five possible approaches that could be employed in this situation, including the pros and cons of each approach.

Methods

Summary of PNRP Research Designs

The PNRP is a collaborative effort involving nine separate research projects (Table 1). Eight of the nine sites contributed data to the national dataset. The ninth site focused solely on the American Indian / Alaska Native population of the northwestern US, and the data sharing agreements are specific to that setting. Each project designed the implementation of its intervention taking into account factors that were unique to their health care delivery system and patient population. Some projects were able to employ true experimental designs, while others implemented variations of quasi-experimental designs.

Four projects – Denver, Ohio, Rochester and Tampa – implemented true experimental designs, using random assignment of either individuals or clinics to the Intervention or Control conditions. Given adequate sample size, random assignment allows the analysis to make the assumption that treatment conditions formed through this process are equivalent, and observed differences between control and intervention treatment conditions can be attributed to the intervention effect.

Denver and Rochester used a more classical RCT design and randomly assigned *individuals* to either a control group or the intervention condition.[6] Ohio and Tampa used a GRT approach and randomly assigned *groups* (i.e., clinics) to either a control group or an intervention group. GRT is a comparative study design in which identifiable groups (i.e., study units) are assigned at random to study conditions and observations are made on members of those groups.[7] Studies with different units of assignment and observation exist in many disciplines and pose a number of design and analytic problems not present when individuals are randomized to study conditions. A central problem is that the intervention effect must be assessed against the between-group variance rather than the within-group variance.[8]

Furthermore, the between-group variance is usually larger if based on identifiable groups than if based on randomly constituted groups. This is the result of the positive intraclass correlation expected among responses from members of the same identifiable group[9]; that correlation reflects an extra component of variation attributable to groups above and beyond that attributable to their members. In addition, the degrees of freedom (df) available to estimate the between-group variance is normally less than that for the within-group variance when there are a limited number of groups per condition. The extra variation and limited df can combine to reduce power and therefore make it difficult to detect important intervention effects in an otherwise well-designed and properly-executed research trial.

Ohio originally identified 12 participating clinics and stratified them according to clinic type – university-based clinic (n = 8) vs. neighborhood health center (n = 4). Within clinic type, individual clinics were rank-ordered on the proportion of African-American patients and pairs were formed by clinics adjacent in rank. Clinics within each pair were then randomly assigned to either Navigation or Control conditions. This method of forming pairs should maintain balance between the two conditions on at least the variable that was used to rank the clinics in the first place. To the extent that this variable is also a surrogate for other variables (e.g., SES, insurance coverage) that might be correlated with the outcome, then this assignment process will permit a more direct interpretation of results. Data on these confounding variables can be examined for balance among the treatment conditions.

Tampa identified 12 clinics nested within five distinct health care organizations (the five organizations having 3, 3, 2, 2, and 2 clinics respectively). There was a further constraint that for health care organizations having 3 clinics, one was randomly assigned to be the Control clinic and the remaining 2 clinics were designated as Navigated clinics. Random assignment of clinics occurred within each health care organization. The Tampa project presents a case where the number of units to be randomly assigned is small and the degree to which equivalency between Control and Navigated conditions was achieved is of less certainty than the situation where several hundred individuals are randomly assigned.

The remaining 5 projects – Boston, Chicago-ACCESS, Chicago-VA, San Antonio and Washington DC, employed a quasi-experimental design,[10] with the distinguishing feature that patients (or clinics) were not randomly assigned to conditions. The decision to use non randomized designs reflected the nature of the intervention and the collaborations with community partners to develop and conduct the research. This leaves attribution of outcomes to Patient Navigation open to other plausible explanations. The same kinds of statistical analyses can be applied to data from quasi-experiments as from true experiments; it is the ability to directly interpret the findings and the confidence in the obtained results that separates these two kinds of designs. Random assignment eliminates many of the alternative plausible explanations for the obtained results that may be more difficult to eliminate in quasi-experimental designs, and facilitates generalizability of results to similar populations.

The Design and Analysis Committee considered the various study designs and proposed five possible analytic methods or approaches for a national evaluation of the major outcomes of the PNRP. Each approach was evaluated by the committee in regards to its strengths, weaknesses, and suitability to address the unique methodological issues of the PNRP. Results of that evaluation are presented below (see also Table 2).

Results

Approach #1: No presentation of combined findings. Each project in the PNRP would be separately analyzed and reported as an independent study

One possible approach to evaluate the PNRP is to make no attempt to combine the data, but instead present each project based on its own analysis. This approach would eliminate the difficulties in assessing the best method to combine data. This approach is useful if the findings of the individual projects are markedly different, allowing readers to more easily see which projects had a significant effect, and which did not. However, this approach does not provide a coherent summary of the overall PNRP program, and leaves the synthesis of the findings to the reader. An understanding of the effects of patient navigation from individual project analyses would only emerge over time as project results were published.

An example of such an analytic approach is the Centers for Medicare and Medicaid Services (CMS) funded program of 15 individual RCTs evaluating care coordination to improve quality of care, re-hospitalization rates and Medicare expenditures for Medicare beneficiaries. In this program, each project developed their own intervention and selected their own target diseases, study population, and established inclusion and exclusion criteria specific to their respective projects.[11] CMS developed a uniform evaluation procedure, using claims data and a standardized telephone survey. To summarize the effects of the overall program, the authors chose to present the data as 15 parallel RCTs, outlining the differences in each intervention and study population.

Interpretation of the findings primarily focused on the two projects that demonstrated positive effects, with an attempt to understand these findings in the context of differing study designs. This approach proved useful in the setting where trials utilized markedly different designs and where the nature of the intervention varied significantly from project to project. In contrast, the PNRP program allowed different study designs, but utilized a standard intervention strategy, study population, inclusion, and exclusion criteria.

Approach #2: Pool the data from all projects and analyze at the individual level ignoring any possible intraclass correlation

This analytical approach would combine data from all the projects, analyze at the individual level and ignore any possible intraclass correlation. The primary analysis would be a test of the mean difference in time from abnormal finding to diagnostic resolution for the Control condition compared to the Intervention (Navigated) condition. The analysis could be performed using standard methods based on the general linear model [12], such as t-tests or F-tests with df based on the number of individuals. While separate analyses could be performed for the different cancer sites – breast, cervical, colorectal, prostate – the basic analytical approach would be the same.

The major advantages of this approach are that it would use all data from all projects in a single analysis and that it would rely on familiar methods based on the general linear model.[12] Pooled data also yield larger sample sizes and increased power, and thereby permit the exploration of rarer events and the examination of the main effects on hypothesized sub-groups, such as patients with co-morbidities.

However, these advantages are offset by several serious disadvantages. A primary disadvantage is that this approach is inappropriate for GRTs and for quasi-experiments involving non-random assignment of identifiable groups. [7, 13] Another is that larger projects would contribute more information than smaller projects; if the results varied systematically by study size, the effects in the larger projects could wash out the effects in the smaller projects.

Approach #3: Pooling data from projects with similar designs

Some of the projects employed similar research designs allowing the possibility of pooling raw data from projects that used similar research designs. For example, two projects employed a GRT design (Ohio and Tampa) that targeted shared cancer sites (breast, colorectal). This approach would analyze data from these two GRT projects, employ statistical methods that are appropriate to the study design, and stratify on project so that overall results would be examined and differences between projects could also be evaluated. An advantage of this situation is that it would allow stratification by cancer site and project simultaneously. The committee considered whether to include data on cervical cancer (collected at Ohio but not Tampa) and concluded that such data could be included, increasing the generalizability of findings. Individual-level data would be analyzed, but for

these two projects, the intraclass correlation expected in the data would be addressed and df would be based on the number of groups, not the number of individuals. The analysis would be based on the general linear mixed model.[14]

Two projects, Denver and Rochester, designed and conducted projects in which individuals within clinics were randomly assigned to either the Control or Navigated conditions. Data from these 2 projects could be pooled for a single analysis using Approach #2. Project (Denver versus Rochester) would be a stratification variable to test the interaction effect. The approach would use standard methods based on the general linear model[12], such as t-tests or F-tests with df based on the number of individuals. An issue complicating this approach in this instance is that the major outcome variables are measuring different time frames; for Denver the time is from abnormal finding to diagnostic resolution in 80% of its population and from cancer diagnosis to initiation of treatment in 20%, while for Rochester the time for most of its patient population is from cancer diagnosis to initiation of treatment and/or completion of primary treatment. Hypothetically, however, it could be possible to pool the data across these two projects since both outcomes are measured in the same unit (# of days). This analytical approach could help answer the general question as to whether patient navigation reduces the time to obtain standard quality cancer care – whether that "care" is diagnostic resolution or initiation of cancer treatment. The challenge would be to find ways to appropriately combine and describe the relative merit of the data in reference to: 1) the overall effects of patient navigation (PN) on quality of cancer care; and 2) specific differential effects of PN on cancer diagnostic resolution and initiation of cancer treatment.

While randomization insures comparability of treatment arms on average, adjustment for patient level covariates is still advisable for several reasons.[15] First, despite randomization, imbalances in patient characteristics can occur by chance. Moreover, unadjusted analyses may yield results that are biased toward the null if there is heterogeneity of risk across strata.[16] Adjusted analyses may provide more precise estimates and a summary result closer to stratum specific results.[16, 17]

The primary advantage of approach three is that it would base the analysis on the design of the project: RCTs would be analyzed as RCTs and GRTs would be analyzed as GRTs. Interpretation would be straightforward in both analyses. The primary disadvantages are that this approach would accommodate only 4 of the 9 projects and present two sets of potentially conflicting results. With each analysis based on only two projects, power would also be reduced relative to other options, and questions may arise regarding unequal sample sizes of projects within type of design.

Approach #4: Prospective Meta-Analysis

Meta-analysis is the statistical synthesis of data from separate but similar projects that are combined so that a quantitative summary of the pooled results can be obtained.[18–21] An extension of this approach is prospective meta-analysis (PMA) in which studies are identified, evaluated and determined to be eligible before the results of any of the studies become known.[22] PMA addresses some of the limitations of retrospective approaches to meta-analysis. For example, retrospective analyses can be influenced by the individual study results, potentially affecting studies that are assessed (publication bias), study selection, what outcomes are assessed, and what treatment and patient subgroups are evaluated. In addition, PMA provides standardization across studies of instruments and variable definitions. PMA is an increasingly utilized approach reported in the literature.[23–32]

Using this PMA approach, data from each project would be analyzed separately using methods appropriate to the project design and an effect size of the intervention (Patient Navigation) would be calculated. These effect sizes would then be combined across projects

using standard meta-analytic techniques to obtain a summary measure of the general effect of Patient Navigation on the timely receipt of standard, quality cancer care.

This approach has several advantages. It recognizes the idiosyncrasies across projects and treats these as random effects. It retains individual projects' research designs and allows for stratification according to research design quality. The PMA approach also avoids another important disadvantage of retrospective meta-analyses, where the researcher conducting the meta-analysis is dependent upon the data that are published in the literature, or must request additional data from the authors. In the PNRP, all original data would be available for analysis.

The data for each project would be analyzed separately, maintaining each project's research design and using statistical methods appropriate to the projects' designs with particular attention to the unit of assignment. Effect sizes could be examined by type of design (Individually randomized, group randomized, non-randomized; or randomized versus non-randomized) to examine differences in the estimated magnitude of the effect of PN as a function of research design type or quality of research design.

A potential disadvantage is the risk that the analyses performed for the PMA might differ from the analyses eventually reported by the individual projects. That risk can be minimized by requiring the individual projects to specify in advance their primary analysis plan. However, if projects plan their primary analysis with the inclusion of data they collected beyond the common dataset, this analysis could not be completely replicated. If the individual projects have very different effect sizes (for example, several with a positive effect of navigation, and several with no effect or a negative effect in the navigated arm) then pooling the effect sizes would not be appropriate and presenting the individual project results would be recommended.

Approach #5: Simulated Group Randomized Design

In this approach, pairs of matched groups would be created within each project to mimic a pair-matched group-randomized trial. The grouping would occur so that the number of observations in each group formed within a particular project is balanced. The matching would be done such that factors most highly associated with outcomes were similarly distributed in both the control and navigated study conditions. Potential matching factors could include cancer site, health insurance coverage, ethnic and racial minority status, gender, and age. The distribution of these factors would first be examined to determine overall balance between navigated and control patients within each individual cancer site. There is no requirement that the number of groups per project be equal or that group as defined for one endpoint be the same group defined for another endpoint.

For projects where group randomization of clinics was used, subjects from each individual clinic would remain together in a single group, and not split into multiple groups even with a large sample size. In those cases where a 'group' was the unit of randomization or assignment, it would not be appropriate to split these units to obtain more groups; this would create groups that were not independent and violate the assumptions of the analysis plan. For some clinics that have small sample sizes it may be necessary to combine across clinics and match on cancer site to attain a sufficient number of observations per group. For projects where individual randomization was used, groups would be created based on cancer site and date of index event, in order to create well matched sets of navigated and control pairs. In the case where reasonable matching within strata has not been achieved with respect to the most influential covariates, it may be necessary to employ a mathematical model in the analysis to control for confounding.

Two methods of analysis could be considered for this approach, ANOVA and simple t-tests (or permutation tests for situations where the assumptions of normality are grossly violated). These methods have been described elsewhere.[7, 33, 34]

Strengths of this approach are that it would utilize all of the data from all eight projects and it would help insure balance in the number of observations in each constructed group within a project and in the number of groups in each condition. Balance is important in group-randomized trials as it helps limit the potential impact of other problems that can occur in GRT data. [35] The size of groups should be similar enough in magnitude that weighting of effects would not be required. Another advantage of this method is that by creating comparable groupings across all PNRP projects, confounding that occurs at the level of groups can be controlled in the analysis. Potential disadvantages are that this is an untested and novel approach, and more importantly, it would not reflect the original design of most of the PNRP projects.

Discussion

The PNRP presents some unusual analytic challenges. The PNRP is clearly distinct from traditional multicenter randomized trials and traditional methods of analysis of multicenter trials could not be applied.[36] Instead, we have considered 5 alternative approaches to the analysis of the PNRP data and examined their strengths and weaknesses. However, as community based participatory research gains strength as a methodology to address health disparities, we anticipate that other research groups will face similar challenges when analyzing multi site studies.

Though we considered Approach #1, we judged the weaknesses to outweigh the strengths. The PNRP was created with the overarching goal of estimating the value of patient navigation. The results from individual projects would certainly add to the research literature on this issue, but a unified summary would not emerge and readers would be left to find, evaluate, and synthesize the reports from the individual projects. In addition, this approach would not allow a thorough evaluation of heterogeneity of navigation effects across projects. The circumstance where this method would however be more applicable is the situation where the effect sizes from the different individual projects are in opposite directions, and therefore a synthesized effect size is not appropriate.

Another major disadvantage to Approach #1 was that a combined analysis should have greater power than a series of project-specific analyses. Pooling data allows sufficient sample size to explore effects of navigation across subpopulations of interest. Moreover, because the number of cancers arising in primary care settings is small, navigation programs that target primary care settings are unlikely to have sufficient numbers of cancers diagnosed to explore effects for each cancer site. By pooling data across cancer sites, there is a greater opportunity to examine the effects of navigation programs that target the full cancer continuum, from diagnosis through treatment.

Two analytic approaches that would pool data cross projects were considered unacceptable by the Design and Analysis Committee. First, simply combining all data and conducting an individual-level analysis would ignore intraclass correlation and each project's unique study design, greatly increasing the likelihood of a type 1 error.[7, 13] In addition, projects having large sample size would be weighted greater in Approach #2 which could unduly influence findings.[37] Approach #3 would combine data only from projects having similar study designs but was deemed undesirable as it would exclude data from five of the eight projects.

Each of these remaining two approaches has potential strengths and weaknesses. Strengths of the PMA approach included an analytic method that was familiar to most readers, the

ability to utilize data from all projects, and the ability to account for unique research designs of each project, the ability to avoid problems common to many meta-analytic efforts by virtue of having complete access to original data, and its prospective approach. A meta-analysis conducted in advance of publication of individual project results could risk inconsistencies with later reports from the separate projects. Inconsistencies can be potentially managed through governance such as enforcing (or at least negotiating) agreed upon analytic methods between that in the pooled analysis and that performed by individual PNRP projects. In addition, differences in analytic methods and resultant discrepancies in reported results between pooled analysis and individual projects would need full disclosure and explanation in subsequent publications.

Advantages of approach #5 which would construct a simulated pair-matched group-randomized trial include utilizing data from all projects, insuring balance between intervention and control conditions, and increased efficiency. Disadvantages include an untested analytic approach not familiar to reviewers or readers, a less transparent approach which lacks adherence to the original designs of the PNRP projects, difficulty achieving sufficient sample size per group, and finally issues of weighting if some projects contribute more groups than others.

Adequate control of confounding is an important analytic issue in all methods. In pooled analyses, confounding can occur at several levels; at the level of the PNRP project and cancer site, at the level of groups that occur within the pooled data (e.g. clinics or hospitals), and at the individual patient level. Factors that are highly correlated are difficult to separate in any analytic strategy. For example, three of the four sites examining cervical cancer are non-randomized while four out of five sites examining colon cancer are randomized. As a result it will be difficult to isolate the potential effects of cancer site from potential effects of the study design. Confounding of group level (e.g. clinic) characteristics is also important in all analytic methods. While group randomized trials have inherent groups that can be assessed, a potential strength of method five is that it creates similar groupings within other study designs allowing for control of confounding at this level.

There were a number of limitations considered in our discussion of analytic approaches. First, the conclusions reached were based on expert opinion and not derived from actual analyses performed. Because data collection is still underway, there was no attempt to perform simulations, for example, that contrasted results from different analytic strategies or assumptions. Finally, the issues faced were unusual and there is little existing literature to put these conclusions into context. Even so, programs that involve different designs and interventions addressing the same problem are not uncommon.

The PNRP faced an unusual situation in which timely program evaluation required the analysis of pooled data arising from projects having heterogeneous study designs. This situation could re-emerge in future multi-site, community based participatory research projects. The PNRP D&A Committee considered several analytic approaches and concluded that a prospective meta-analysis is one appropriate analytic strategy in these situations. A novel simulated group randomized approach was also proposed as an alternative analytic approach. Future research (e.g. data simulations) would help to understand how program evaluation results are influenced by the analytic method used.

Acknowledgments

The Patient Navigation Research Program is supported by NIH Grants U01 CA116892, U01 CA 117281, U01CA116903, U01CA116937, U01CA116924, U01CA116885, U01CA116875, U01CA116925, American Cancer Society #SIRSG-05-253-01. The contents are solely the responsibility of the authors and do not necessarily

represent the official views of the Center to Reduce Cancer Health Disparities, NCI/NIH or the American Cancer Society.

Abbreviations

PNRP	Patient Navigation Research Program
NCI	National Cancer Institute's
CRHCD	Center to Reduce Cancer Health Disparities
RCT	randomized clinical trials
GRT	group-randomized trials
QE	quasi-experimental designs

References

1. Jean-Pierre P, Hendren S, Fiscella K, et al. Understanding the Processes of Patient Navigation to Reduce Disparities in Cancer Care: Perspectives of Trained Navigators from the Field. J Cancer Educ. 2010 [Epub ahead of print].
2. Wells KJ, Battaglia TA, Dudley DJ, et al. Patient navigation: state of the art or is it science? Cancer. Oct 15; 2008 113(8):1999–2010. [PubMed: 18780320]
3. Robinson-White S, Conroy B, Slavish KH, Rosenzweig M. Patient navigation in breast cancer: a systematic review. Cancer Nurs. Mar-Apr;2010 33(2):127–140. [PubMed: 20142736]
4. Friedenreich CM. Methods for pooled analyses of epidemiologic studies. Epidemiology. Jul; 1993 4(4):295–302. [PubMed: 8347739]
5. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int J Epidemiol. Feb; 1999 28(1):1–9. [PubMed: 10195657]
6. Friedman, L.; Furberg, C.; DeMets, D. Fundamentals of Clinical Trials. 3. New York, New York: Springer-Verlag, Inc; 1998.
7. Murray, D. Design and Analysis of Group-Randomized Trials. New York, NY: Oxford University Press; 1998.
8. Cornfield J. Randomization by group: a formal analysis. Am J Epidemiol. Aug; 1978 108(2):100–102. [PubMed: 707470]
9. Kish, L. Survey Sampling. New York, NY: John Wiley & Sons; 1965.
10. Shadish, W.; Cook, T.; Campbell, D. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin Company; 2002.
11. Peikes D, Chen A, Schore J, Brown R. Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials. Jama. Feb 11; 2009 301(6):603–618. [PubMed: 19211468]
12. Searle, S. Linear Models. New York, NY: John Wiley & Sons; 1971.
13. Donner, A.; Klar, N. Design and Analysis of Cluster Randomization Trials in Health Research. London: Arnold; 2000.
14. Harville DA. Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. J American Statistical Association. 1977; 72(358):320–338.
15. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Bmj. 2010; 340:c869. [PubMed: 20332511]
16. Kent DM, Trikalinos TA, Hill MD. Are unadjusted analyses of clinical trials inappropriately biased toward the null? Stroke. Mar; 2009 40(3):672–673. [PubMed: 19164784]
17. Steyerberg EW, Bossuyt PM, Lee KL. Clinical trials in acute myocardial infarction: should we adjust for baseline characteristics? Am Heart J. May; 2000 139(5):745–751. [PubMed: 10783203]

18. Glass, G.; McGaw, B.; Smith, M. Meta-analysis in Social Research. Beverly Hills, CA: Sage Publications; 1981.
19. Hedges, L.; Olkin, I. Statistical Methods for Meta-Analysis. New York, NY: Academic Press; 1985.
20. Borenstein, M.; Hedges, L.; Higgins, J.; Rothstein, H. Introduction to Meta-Analysis. West Sussex, UK: John Wiley & Sons, Ltd; 2009.
21. Grissom, R.; Kim, J. Effect Sizes for Research: A Broad Practical Approach. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2005.
22. [Accessed August 8, 2011.] Cochrane Prospective Meta-Analysis Methods Group. http://pma.cochrane.org
23. World Health Organization-International Society of Hypertension Blood Pressure Lowering Treatment Trialists' Collaboration. Protocol for prospective collaborative overviews of major randomized trials of blood-pressure-lowering treatments. J Hypertens. Feb; 1998 16(2):127–137. [PubMed: 9535138]
24. Margitic SE, Morgan TM, Sager MA, Furberg CD. Lessons learned from a prospective meta-analysis. J Am Geriatr Soc. Apr; 1995 43(4):435–439. [PubMed: 7706636]
25. Turok DK, Espey E, Edelman AB, et al. The methodology for developing a prospective meta-analysis in the family planning community. Trials. 2011; 12:104. [PubMed: 21527040]
26. Cholesterol Treatment Trialists' (CTT) Collaboration. Protocol for a prospective collaborative overview of all current and planned randomized trials of cholesterol treatment regimens. Cholesterol Treatment Trialists' (CTT) Collaboration. Am J Cardiol. Jun 1; 1995 75(16):1130–1134. [PubMed: 7762499]
27. Margitic SE, Inouye SK, Thomas JL, Cassel CK, Regenstreif DI, Kowal J. Hospital Outcomes Project for the Elderly (HOPE): rationale and design for a prospective pooled analysis. J Am Geriatr Soc. Mar; 1993 41(3):258–267. [PubMed: 8440849]
28. Baigent C, Keech A, Kearney PM, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet. Oct 8; 2005 366(9493):1267–1278. [PubMed: 16214597]
29. Valsecchi MG, Masera G. A new challenge in clinical research in childhood ALL: the prospective meta-analysis strategy for intergroup collaboration. Ann Oncol. Dec; 1996 7(10):1005–1008. [PubMed: 9037357]
30. Askie LM, Baur LA, Campbell K, et al. The Early Prevention of Obesity in CHildren (EPOCH) Collaboration--an individual patient data prospective meta-analysis. BMC Public Health. 2010; 10:728. [PubMed: 21106110]
31. Design rational, and baseline characteristics of the Prospective Pravastatin Pooling (PPP) project--a combined analysis of three large-scale randomized trials: Long-term Intervention with Pravastatin in Ischemic Disease (LIPID), Cholesterol and Recurrent Events (CARE), and West of Scotland Coronary Prevention Study (WOSCOPS). Am J Cardiol. Nov 1; 1995 76(12):899–905. [PubMed: 7484829]
32. Province MA, Hadley EC, Hornbrook MC, et al. The effects of exercise on falls in elderly patients. A preplanned meta-analysis of the FICSIT Trials. Frailty and Injuries: Cooperative Studies of Intervention Techniques. Jama. May 3; 1995 273(17):1341–1347. [PubMed: 7715058]
33. Donner A. Statistical methodology for paired cluster designs. Am J Epidemiol. Nov; 1987 126(5): 972–979. [PubMed: 3661544]
34. Donner A, Donald A. Analysis of data arising from a stratified design with the cluster as unit of randomization. Stat Med. Jan-Feb;1987 6(1):43–52. [PubMed: 3576016]
35. Gail MH, Mark SD, Carroll RJ, Green SB, Pee D. On design considerations and randomization-based inference for community intervention trials. Stat Med. Jun 15; 1996 15(11):1069–1092. [PubMed: 8804140]
36. Worthington H. Methods for pooling results from multi-center studies. J Dent Res. 2004; 83(Spec No C):C119–121. [PubMed: 15286137]
37. Bravata DM, Olkin I. Simple pooling versus combining in meta-analysis. Eval Health Prof. Jun; 2001 24(2):218–230. [PubMed: 11523387]

Patient Navigation Research Program Investigators

Program Office

National Cancer Institute, Center to Reduce Cancer Health Disparities: Martha Hare, Mollie Howerton, Ken Chu, Emmanuel Taylor, Mary Ann Van Duyn.

Evaluation Contractor

NOVA Research Company: Paul Young, Frederick Snyder.

Clinical Centers

Boston University Women's Interdisciplinary Research Center: PI-Karen L. Freund, Tracy A. Battaglia.

Denver Health and Hospital Authority: PI-Peter Raich, Elizabeth Whitley.

George Washington University Cancer Institute: PI-Steven R. Patierno, Lisa M. Alexander, Paul H. Levine, Heather A. Young, Heather J. Hoffman, Nancy L. LaVerda.

H. Lee Moffitt Cancer Center and Research Institute: PI-Richard G. Roetzheim, Cathy Meade, Kristen J. Wells.

Northwest Portland Area Indian Health Board: PI- Victoria Warren-Mears.

Northwestern University Robert H. Lurie Comprehensive Cancer Center: PI-Steven Rosen, Melissa Simon.

Ohio State University: PI-Electra Paskett.

University of Illinois at Chicago and Access Community Health Center: PI-Elizabeth Calhoun, Julie Darnell.

University of Rochester: PI-Kevin Fiscella, Samantha Hendren.

University of Texas Health Science Center at San Antonio Cancer Therapy and Research Center: PI-Donald Dudley, Kevin Hall, Anand Karnard, Amelie Ramirez.

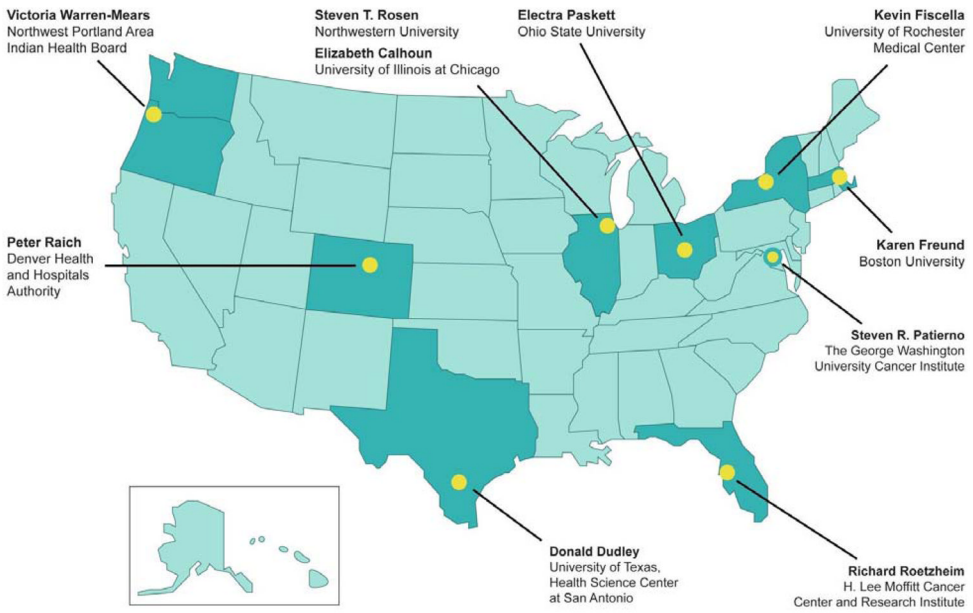


Figure 1.
Location of PNRP Projects

Table 1

Design Characteristics of PNRP Projects

	Boston	Chi-Access	Chi-VA	Denver	Ohio	Rochester	San Ant	Tampa	Wash-DC
Targeted Cancer Sites									
Breast	X	X		X	X	X	X	X	X
Cervical	X	X			X		X		
Colorectal			X	X	X	X		X	
Prostate			X	X					
Design									
Randomized – Individual level				X		X			
Randomized – Group level					X			X	
Non-Randomized	X	X	X				X		X
Unit of Assignment									
Individual				X		X			
Group	X	X	X		X		X	X	X
Controls									
Actively Recruited Controls				X	X	X			
Medical Records-Based Controls	X	X	X				X	X	X
Clinics/Care sites Number (#Nav/#Ctl)	6 (3/3)	9 (5/4)	1	1	16 (8/8)	11	5 (2/3)	12 (7/5)	15 (8/7)
Intervention Scope									
Abnormality to Diagnosis	X	X	X	X	X	X	X	X	X
Diagnosis to Start of Treatment	X		X	X	X	X	X	X	X
Start to Completion of Treatment	X		X	X	X	X	X	X	X
Enrollment as of 6/1/11									
Total subjects	**3042**	**1023**	**513**	**1249**	**674**	**344**	**1052**	**1320**	**2594**
Subjects with cancer	**196**	**118**	**0**	**171**	**27**	**321**	**9**	**53**	**349**
Total PNRP Enrollment (all projects)									
Total subjects **11,643**									
Total subjects with cancer **1,048**									

Table 2

Comparison of Potential Analysis Methods

	Analysis Method	Description	PROs	CONs
Approach #1	Analyze and report each project separately	Each PNRP project would be separately analyzed and reported as an independent study. There would be no attempt to combine data from different projects.	• Accounts for differing study designs	• No summary measure of effect • Leaves synthesis of findings to readers • Summary of overall program may be biased by projects with favorable results.
Approach #2	Individual-level analysis ignoring any intraclass correlation	Combine data from all the projects, analyze using standard methods, ignoring any intraclass correlation All individuals receive equal weight in the analysis. Primary analysis compares mean difference in time to diagnostic resolution for navigated versus control subjects.	• All data are utilized – no subjects excluded	• Pooling data ignores each project's internal validity • Ignores "group" in group-randomized designs and non-randomized group designs • Projects with large sample sizes could overwhelm results
Approach #3	Pool data from projects having similar designs	Data from two projects that used group-randomized designs would be combined and analyzed. Two projects that randomized at individual level would similarly be pooled and analyzed.	• Interpretation more straightforward • Restricts analysis to projects with true experimental designs • Increases generalizability of findings	• Only four of the nine projects contribute to the analysis • Analyses yield two separate results
Approach #4	Prospective meta analysis	Data from each project analyzed separately to estimate effect size. Effect sizes would be combined using meta-analytic techniques to obtain overall measure of program effect.	• All data are available, not just that in published literature • Measures of effect for each program are computed • Allows comparisons across projects (by design, by type of navigation) • Avoids publication bias	• Meta-analysis in advance of the publication of project results risks inconsistencies with the reports by the projects
Approach #5	Simulated Group Randomized Design	For projects that did not randomize by group, pairs of matched, and balanced groups will be created from the individual data simulating a group randomized design.	• Utilizes all data from nine projects • Insures balance between intervention and control arms • Allows for control of confounding at level of 'group' across projects • Efficient use of available data	• Untested and novel approach. • Less transparent approach • Does not reflect original study designs in individual projects. • Projects with more groups have greater influence on analysis Small number of observations per group may require combining groups

Editorial

Improving the Management of Warfarin May Be Easier Than We Think

Adam J. Rose, MD, MSc, FACP

Performance variation is easy to find—all one needs to do is to look for it. We have long known that some nations achieve better control of hypertension than others,[1] some hospitals have shorter door-to-balloon times than others,[2] and some cardiac surgeons have better risk-adjusted mortality after coronary artery bypass graft surgery than others.[3] In fact, it is difficult to recall an instance when performance was found not to vary. Given that performance variation is ubiquitous, it is no longer shocking to find it; the more interesting question is why performance varies so much. The answer to this question would likely be a key step along the pathway to improving performance.

Article see p 2309

It is rather uncommon to find a single answer to the question of why performance variation exists. However, in the current issue of *Circulation*, Van Spall and colleagues[4] have found an unusually straightforward explanation for performance variation, at least in the context of the management of warfarin. The authors found that site-level adherence to a relatively simple algorithm regarding when to change the dose of warfarin and when not to change the dose predicted fully 87% of between-center variance. Adding patient-level clinical variables (ie, risk-adjustment), center-level variables, and country-level variables only increased the amount of explained variation to 89%. In short, management of warfarin doses appeared to be almost deterministic regarding the anticoagulation control that was achieved. Remarkably, the authors also found that greater adherence to the algorithm also predicted a reduced rate of the combined primary end point of stroke, major hemorrhage, or death at the site level. For each 10% increase in algorithm-consistent dosing at the center level, the annual rate of the combined end point was 8% lower, even after adjusting for a host of patient-level predictors. The algorithm has additional attractive features as well, not least of which is that it does not require a computer or proprietary software to use. In fact, it would be equally suitable to use in developing countries.

The authors deftly used a feature of their dataset to demonstrate that percent time in therapeutic range (TTR) really is in the causal pathway to adverse events. Because their dataset was drawn from a randomized trial of dabigatran versus warfarin, they were able to show that, whereas sites with more algorithm-consistent warfarin dosing had lower rates of adverse events among patients receiving warfarin, they did not have lower rates of adverse events among patients receiving dabigatran. Some have expressed doubts about whether TTR really is in the causal pathway to outcomes,[5] suggesting that instead, sites with higher TTR also might be delivering high-quality care in other ways. If this were true, then quality improvement efforts aimed at increasing TTR might not achieve the desired benefits in terms of preventing adverse events. The finding by Van Spall and colleagues[4] serves as a strong refutation for this line of reasoning, because if high-TTR sites were truly delivering other interventions responsible for preventing adverse events, we would have seen a similar reduction in adverse events among patients receiving dabigatran. The present study therefore serves as a strong endorsement of efforts to improve TTR at the site level and thereby prevent adverse events.

The algorithm studied by Van Spall and colleagues[4] was similar to others that are in widespread use. The algorithm suggested no dose change for an in-range international normalized ratio (INR) value, a 10% dose change when the INR was somewhat out of range (1.51–1.99 or 3.01–4.00), and a 15% dose change for greater deviations from the target range. The authors note that, although they distributed this algorithm, they cannot know to what extent the sites of care actually used it—in essence, their study examined the difference between care that would have been algorithm-concordant, as the authors put it, "whether intentionally or not." In a study by Kim et al,[6] a very similar (but not completely identical) dosing algorithm was shown to improve TTR. The Kim study complements the current study in that it shows that the introduction of a standardized dosing algorithm improved TTR considerably using a before–after design.[6] Although Van Spall and colleagues were able to show that algorithm-concordant care was important, they were unable to examine whether the introduction of the algorithm had actually increased the likelihood that such care would be delivered. Also, in the Kim study, algorithm-consistent dosing was highly correlated with patient-level TTR, which complements the site-level findings of the current study. Taken together, the 2 studies suggest that the algorithm is both valid (in the sense of predicting important outcomes)[4] and feasible (in the sense that it can successfully be introduced and change clinical practice).[6]

From the Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, Bedford, MA; and the Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA.

Correspondence to Adam J. Rose, MD, MSc, FACP, Center for Health Quality, Outcomes, and Economic Research, 200 Springs Road, Mail Stop 152, Bedford, MA 01730. E-mail adamrose@bu.edu

(*Circulation*. 2012;126:2277-2279.)

***Circulation* is available at http://circ.ahajournals.org**
DOI: 10.1161/CIRCULATIONAHA.112.141887

Table. Some Key Challenges to Improving Anticoagulation Control

Topic	Issue	Challenge or Knowledge Gap
Warfarin dose management	Limited use of computerized dose support or paper-based algorithms, both of which are shown to improve control	Promoting wider adoption of computerized dose support or paper-based algorithms
Loss to follow-up	Lack of functioning systems to track patients and detect loss to follow-up	Need to develop such systems, particularly for sites without proprietary anticoagulation management software
Use of nonstandard target INR ranges	Lower target INR ranges such as 1.5 to 2.0 remain in use, despite having been shown to produce harms without additional benefit	Finding effective ways to promote the use of standard rather than nonstandard target ranges
Timely follow-up after deranged INR values	Prompt follow-up after high (>4) or low (≤1.5) INR improves control	Clinicians may resist changing practice; patients may resist frequent follow-up
Patient education	Patient education may lack standardization and may have limited effectiveness.	Developing new and innovative ways to educate and re-educate patients about warfarin therapy
Patient adherence to therapy	Limited adherence to pill-taking, dietary consistency, on-time follow-up, and other matters	Need to identify and address various patient-level barriers to improved adherence
Anticoagulation clinic leadership	Each clinic needs strong leadership to ensure continuous quality improvement	Individuals may not feel empowered to identify and implement strategies to improve outcomes
Performance measurement	Anticoagulation therapy is clearly important enough to deserve a program of performance measurement	Performance measures have been developed, but still need to be used more widely, and may need to be amended over time in response to feedback

INR indicates international normalized ratio.

Any algorithm for warfarin dose management, whether paper-based or computer-based, needs to be introduced with a caveat, namely that the clinician must always have the power to override the algorithm. Any decision about warfarin dose changes must occur in the context of the visit, and some information that is divulged by the patient (for example, recent dietary intake) may well prompt a decision that is discordant with the algorithm.[7] However, both the Van Spall and Kim studies strongly suggest that clinicians managing warfarin should ask themselves whether they really have a compelling reason to deviate from the algorithm—and then ask again for good measure.

The present study also adds to a growing discussion about how to manage warfarin doses when the INR is only slightly out of range. In the past, several studies have suggested that dose changes are not necessary for mildly deranged INR, and in fact may begin a cycle of overcorrection and rebound that may worsen control.[8–9] Based on this limited evidence, the latest consensus guidelines for managing warfarin suggest not changing the dose when the INR deviates by ≤0.5 from the target range, and instead merely rechecking in 1 to 2 weeks.[10] Notably, a deviation of up to 0.5 is an even wider tolerance range than would be suggested by either of the previous studies.[8–9] With this issue in mind, Van Spall and colleagues examined whether their findings would also be true when the tolerance range was widened from 2.0 to 3.0 to 1.9 to 3.1, 1.8 to 3.2, 1.7 to 3.3, and 1.6 to 3.4. They found that TTR was higher with increased adherence to any of these ranges, but was best with the strict construction of the target range (ie, 2.0–3.0). In addition, only the strict construction predicted the primary combined end point in a statistically significant fashion. Thus, the findings of the present study cast at least some doubt on the idea of a widened tolerance range, within which slightly deranged INR values do not prompt a dose change. Ideally, a well-designed prospective study could address this issue more definitively. For now, although some clinicians may still opt to have a tolerance range, it may be prudent not to extend it beyond approximately 1.8 to 3.2,[8–9] as opposed to the much wider range of 1.5 to 3.5 recommended by the most recent guidelines.[10]

One noteworthy limitation of the present study stems from its origins in a clinical trial.[11] The patients represented in this study may have been unusually adherent to their medication, and gaps in INR monitoring were likely minimal. In a real-world setting, dose management may not predict 87% of site variation, and other issues, like gaps in INR monitoring, also may play an important role. Our group has shown that nonstandard target ranges, gaps in INR monitoring, and failure to recheck INR promptly after an out-of-range value are also strong predictors of site-level TTR.[12–14] Together, these 3 measures predicted 48% of TTR variation among sites of care in the Veterans Health Administration (our unpublished finding). It is likely, therefore, that although adherence to the algorithm examined by Van Spall and colleagues may be the single most important determinant of site-level TTR, it is not the only performance measure worth addressing. In fact, improving TTR is likely to require a multifactorial approach, as there are many related issues that contribute to TTR (Table).

Another limitation, mentioned above, is that the present study was not an attempt to actually implement the algorithm under study. Kim and colleagues successfully implemented a similar algorithm,[6] but they did so at a single site of care, which may not be representative of other sites. An important next step would be a prospective test of this algorithm in a real-world setting, preferably across an integrated health system. Such an effort would require great attention to the principles of implementation science to promote successful adoption of the algorithm.[15] In fact, we have recently begun such a study in the New England region of the Veterans

Health Administration. One challenge for our effort, and for any effort occurring in the real world, is that automated data may not reliably identify dose changes in warfarin, which is usually prescribed as use as directed. A need to rely on chart review to track adherence to the algorithm may hamper efforts to measure its uptake.

In summary, Van Spall and colleagues[4] have added several important new pieces of information to the literature. First, in combination with other studies,[6] the present study convincingly demonstrates that managing warfarin according to a standard algorithm can improve patient outcomes. Second, it should lay to rest any lingering doubts about whether TTR is truly in the causal pathway to adverse events. Finally, it suggests that changing the dose of warfarin for any out-of-range INR may in fact be best, although other studies have found differently,[8–9] and further evidence is needed. The study by Van Spall and colleagues certainly provides a powerful argument for greater adoption of any sort of system to promote standardized dose management for warfarin—either computer support when it is available, or paper-based algorithms if not. Improving outcomes for patients receiving warfarin will not be easy, but it may be easier than we think.

Disclosures

Dr Rose is supported by a career development award from the Veterans Health Administration, Health Services Research and Development Service (CDA-2-08-017). The opinions expressed are those of the author and do not necessarily represent the official views or policies of the Veterans Health Administration.

References

1. Wolf-Maier K, Cooper RS, Banegas JR, Giampaoli S, Hense HW, Joffres M, Kastarinen M, Poulter N, Primatesta P, Rodriguez-Artalejo F, Stegmayr B, Thamm M, Tuomilehto J, Vanuzzo D, Vescio F. Hypertension prevalence and blood pressure levels in 6 European countries, Canada, and the United States. *JAMA*. 2003;289:2363–2369.
2. Bradley EH, Herrin J, Wang Y, McNamara RL, Radford MJ, Magid DJ, Canto JG, Blaney M, Krumholz HM. Door-to-drug and door-to-balloon times: where can we improve? Time to reperfusion therapy in patients with ST-segment elevation myocardial infarction (STEMI). *Am Heart J*. 2006;151:1281–1287.
3. Jha AK, Epstein AM. The predictive accuracy of the New York State coronary artery bypass surgery report-card system. *Health Aff (Millwood)*. 2006;25:844–855.
4. Van Spall HGC, Wallentin L, Yusuf S, Eikelboom JW, Nieuwlaat R, Yang S, Kabali C, Reilly PA, Ezekowitz MD, Connolly SJ. Variation in warfarin dose adjustment practice is responsible for differences in the quality of anticoagulation control between centers and countries: an analysis of patients receiving warfarin in the Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) trial. *Circulation*. 2012;126:2309–2316.
5. van Walraven C, Oake N, Coyle D, Taljaard M, Forster AJ. Changes in surrogate outcomes can be translated into clinical outcomes using a Monte Carlo model. *J Clin Epidemiol*. 2009;62:1306–1315.
6. Kim YK, Nieuwlaat R, Connolly SJ, Schulman S, Meijer K, Raju N, Kaatz S, Eikelboom JW. Effect of a simple two-step warfarin dosing algorithm on anticoagulant control as measured by time in therapeutic range: a pilot study. *J Thromb Haemost*. 2010;8:101–106.
7. Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest*. 2008;133:160S–198S.
8. Banet GA, Waterman AD, Milligan PE, Gatchel SK, Gage BF. Warfarin dose reduction vs watchful waiting for mild elevations in the international normalized ratio. *Chest*. 2003;123:499–503.
9. Rose AJ, Ozonoff A, Berlowitz DR, Henault LE, Hylek EM. Warfarin dose management affects INR control. *J Thromb Haemost*. 2009;7:94–101.
10. Holbrook A, Schulman S, Witt DM, Vandvik PO, Fish J, Kovacs MJ, Svensson PJ, Veenstra DL, Crowther M, Guyatt GH. Evidence-based management of anticoagulant therapy: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141:e152S–e184S.
11. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener HC, Joyner CD, Wallentin L. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med*. 2009;361:1139–1151.
12. Rose AJ, Berlowitz DR, Miller DR, Hylek EM, Ozonoff A, Zhao S, Reisman JI, Ash AS. INR targets and site-level anticoagulation control: results from the Veterans AffaiRs Study to Improve Anticoagulation (VARIA). *J Thromb Haemost*. 2012;10:590–595.
13. Rose AJ, Hylek EM, Berlowitz DR, Ash AS, Reisman JI, Ozonoff A. Prompt repeat testing after out-of-range INR values: a quality indicator for anticoagulation care. *Circ Cardiovasc Qual Outcomes*. 2011;4:276–282.
14. Rose AJ, Miller DR, Ozonoff A, Berlowitz DR, Ash AS, Zhao S, Reisman JI, Hylek EM. Gaps in monitoring during oral anticoagulation: insights into care transitions, monitoring barriers, and medication nonadherence. *Chest*. In press.
15. Stetler CB, Damschroder LJ, Helfrich CD, Hagedorn HJ. A guide for applying a revised version of the PARIHS framework for implementation. *Implement Sci*. 2011;6:99.

KEY WORDS: Editorials ■ anticoagulants ■ medication therapy management ■ outcome assessment (health care) ■ quality of health care

Journal of Thrombosis and Haemostasis, **10**: 590–595 DOI: 10.1111/j.1538-7836.2012.04649.x

ORIGINAL ARTICLE

INR targets and site-level anticoagulation control: results from the Veterans AffaiRs Study to Improve Anticoagulation (VARIA)

A. J. ROSE,*†, D. R. BERLOWITZ,*†‡, D. R. MILLER,*‡, E. M. HYLEK,*†, A. OZONOFF,*§ S. ZHAO,* J. I. REISMAN* and A. S. ASH*†¶

**Center for Health Quality, Outcomes and Economic Research, Bedford VA Medical Center, Bedford, MA; †Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA; ‡Department of Health Policy and Management, Boston University School of Public Health, Boston, MA; §Biostatistics Section, Boston Children's Hospital, Boston, MA; and ¶Department of Quantitative Health Sciences, Division of Biostatistics and Health Services Research, University of Massachusetts School of Medicine, Worcester, MA, USA*

To cite this article: Rose AJ, Berlowitz DR, Miller DR, Hylek EM, Ozonoff A, Zhao S, Reisman JI, Ash AS. INR targets and site-level anticoagulation control: results from the Veterans AffaiRs Study to Improve Anticoagulation (VARIA). *J Thromb Haemost* 2012; **10:** 590–5.

Summary. *Background:* Not all clinicians target the same International Normalized Ratio (INR) for patients with a guideline-recommended target range of 2–3. A patient's mean INR value suggests the INR that was actually targeted. We hypothesized that sites would vary by mean INR, and that sites of care with mean values nearest to 2.5 would achieve better anticoagulation control, as measured by per cent time in therapeutic range (TTR). *Objectives:* To examine variations among sites in mean INR and the relationship with anticoagulation control in an integrated system of care. *Patients/Methods:* We studied 103 897 patients receiving oral anticoagulation with an expected INR target between 2 and 3 at 100 Veterans Health Administration (VA) sites from 1 October 2006 to 30 September 2008. Key site-level variables were: proportion near 2.5 (that is, percentage of patients with mean INR between 2.3 and 2.7) and mean risk-adjusted TTR. *Results:* Site mean INR ranged from 2.22 to 2.89; proportion near 2.5, from 30 to 64%. Sites' proportions of patients near 2.5, below 2.3 and above 2.7 were consistent from year to year. A 10 percentage point increase in the proportion near 2.5 predicted a 3.8 percentage point increase in risk-adjusted TTR ($P < 0.001$). *Conclusions:* Proportion of patients with mean INR near 2.5 is a site-level 'signature' of care and an implicit measure of targeted INR. This proportion varies by site and is strongly associated with site-level TTR. Our study suggests that sites wishing to improve TTR, and thereby improve patient outcomes, should avoid the explicit or implicit pursuit of non-standard INR targets.

Correspondence: Adam J. Rose, Center for Health Quality, Outcomes and Economic Research, Bedford VA Medical Center, 200 Springs Road, Building 70, Bedford, MA 01730, USA.
Tel.: +1 781 687 3198; fax: +1 917 591 3104.
E-mail: adamrose@bu.edu

Received 20 November 2011, accepted 26 January 2012

Keywords: ambulatory care, anticoagulants, medication therapy management, quality of healthcare, warfarin.

Millions of patients receive warfarin each year to prevent or treat thromboembolic disease. Better anticoagulation control (i.e. a greater percentage of time in therapeutic range [TTR]), can reduce the occurrence of adverse events [1–4]. Therefore, to improve patient outcomes, sites of care should develop systems to measure and improve TTR [5–7]. Ideally, we should measure not only intermediate outcomes (such as TTR), but also processes of care, because process deficiencies provide a prescription for remediation. Because the most useful process measures are 'tightly linked' to outcomes [8], it would be ideal to demonstrate which processes of care are associated with better anticoagulation control.

We have profiled 100 sites of care in an integrated healthcare system (the Veterans Health Administration, or VA) on TTR [9]. Having profiled sites on anticoagulation control (an intermediate outcome of care), we sought to determine which processes of care predict this outcome [8]. Several site-level processes of care are related to TTR, most notably follow-up intervals after out-of-range International Normalized Ratio (INR) values [10]. Another process measure likely to affect anticoagulation control is pursuit of a target INR range of 2–3 for most patients. Several high-quality randomized trials have demonstrated that aiming for a lower INR target range (such as 1.5–2) produces inferior protection from thromboembolism with no reduction in rates of major hemorrhage [11–16]. Despite the established value of the standard target range of 2–3 for patients with atrial fibrillation (AF) or venous thromboembolism (VTE) [5,17,18], some clinicians may continue to implicitly or explicitly aim for non-standard target ranges. Although these clinicians may believe that they can successfully keep patients within a very narrow target range such as 2–2.5, target ranges narrower than a full INR unit do not reduce the variability of INR [19].

One prominent study includes use of a standard target range of 2–3 for patients with AF among its quality measures [20]. However, in most large automated databases, we do not have direct access to the target range, complicating efforts to use this as a quality measure. In this study, we propose a new process of care measure for oral anticoagulation: the proportion of patients at a site, anticoagulated for AF or VTE, who achieve a mean INR of 2.3–2.7. We assert that this does not merely measure differences in INR variability among sites of care, but actually measures site-level propensity to target non-standard target ranges, because even patients with highly variable INR values will still tend to achieve a mean INR very close to what is being sought. We used a large database from the VA to address three main questions. (i) At which level of mean INR do patients record the highest TTR? (ii) Do sites of care vary in their propensity to aim for non-standard target ranges? (iii) Are these differences associated with site-level anticoagulation control? Demonstration of variability of process and a relationship with an intermediate outcome of care would provide strong support for measuring, and intervening to optimize, this new process measure.

Methods

Data

The database for this study has also been described elsewhere [9]. The Veterans AffaiRs Study to Improve Anticoagulation (VARIA) included all patients deemed to be receiving oral anticoagulation therapy from the VA between 1 October 2006 and 30 September 2008, based on the criteria described below. The study was approved by the Institutional Review Board of the Bedford VA Medical Center.

Patients

We included all patients who received warfarin from the VA during the 2-year study period (i.e. at least 30 days' worth dispensed by the pharmacy). We excluded patients whose primary indication to receive warfarin was valvular heart disease. Many such patients have a target INR range of 2.5–3.5 rather than the more standard 2–3, but it is not possible to determine with certainty which patients have the higher target range. Without specific knowledge of the target range, we cannot calculate TTR. For this study, we also excluded the inception period, or the first 6 months of warfarin therapy, a period when many patients may have erratic INR control and may spend much time below the target range.

Laboratory values and calculation of per cent time in range (TTR)

We included INR values within the VA system during times when a patient was 'on warfarin': that is, when a patient was either (i) in possession of warfarin or (ii) having INR tests at least every 42 days. We defined the period of warfarin possession as the duration of the most recent VA prescription for warfarin, plus 30 days. Because patients may be instructed to take half-doses of warfarin, we recognize that going more than 30 days beyond the end of a prescription does not necessarily indicate that warfarin therapy has stopped. We therefore also allowed a consistent pattern of INR measurements (i.e. every 42 days or less) to indicate that a patient was still being managed.

We excluded INR tests measured while the patient was hospitalized. Patients who are hospitalized may receive temporary parenteral anticoagulation or no anticoagulation, so low INR values while hospitalized may be intentional. We calculated TTR using Rosendaal's method [21], which uses linear interpolation to assign an INR value to each day between successive observed INR values. Gaps of 56 days or more between INR values are not interpolated. After interpolation, the percentage of time during which the interpolated INR values lie between 2.0 and 3.0 (from 0 to 100%) is calculated [21].

Sites of care

We included 100 VA sites of care, each of which includes a hospital, an outpatient care center and several outlying community-based clinics. Each site has a specialized anticoagulation clinic, which is usually run by clinical pharmacists under the supervision of a medical director[22]. Therefore, essentially all patients whose anticoagulation is managed in the VA are managed by specialized anticoagulation clinics. Most patients only visited one site of care, and their INR data were assigned to that site. If a patient visited more than one site (3% of patients), we partitioned their data by site.

While most INR values in the VA are processed via an automated laboratory, a few VA sites mostly or exclusively rely upon point-of-care (POC) testing. Our data do not specify which INR tests were obtained via POC. Home testing with POC devices is not covered by the VA, and therefore is extremely uncommon.

Risk adjustment model

We have previously described our risk adjustment model for TTR [9]. We considered many potential variables for inclusion in the model which were likely to impact TTR, including demographics, area-level poverty, driving distance to care, physical health conditions, mental health conditions, number of medications, and number of hospitalizations. Most variables were retained within the model, with the exception of several co-morbid conditions with small effect sizes. The model was derived and validated according to customary procedures, which included considerations of maximizing predictive ability, clinical credibility, and ease of use and understanding. This patient-level risk adjustment model for TTR has an R^2 of 13.3% when used with this dataset [9]. Table 1 contains all the variables in the model.

Table 1 Patient characteristics (n = 103 897)

Variable	Number (%), median (IQR) or mean (SD)
Female gender	1976 (1.9)
Median age (IQR)	72 (62–79)
Race/ethnicity	
Non-Hispanic white	88481 (85.2)
Non-Hispanic black	9572 (9.2)
Hispanic	3229 (3.1)
Asian	392 (0.4)
Native American	449 (0.4)
Other/unknown	1774 (1.7)
Median per cent poverty in zip code of residence (IQR)	10.7 (6.6–16.0)
Median distance from nearest VA facility in miles (IQR)	7.8 (3.7–16.5)
Primary indication for warfarin*	
Atrial fibrillation	66 756 (64.3)
Venous thromboembolism	28350 (27.3)
All others combined	8791 (8.5)
Physical co-morbid conditions	
Cancer (newly diagnosed)	7033 (6.8)
Chronic kidney disease	14581 (14.0)
Chronic liver disease	1228 (1.2)
Chronic lung disease	30 311 (29.2)
Diabetes	41 507 (40.0)
Epilepsy	2903 (2.8)
Heart failure	33 727 (32.5)
Hyperlipidemia	78 205 (75.3)
Hypertension	87 138 (83.9)
Mental co-morbid conditions	
Alcohol abuse	9583 (9.2)
Bipolar disorder	2359 (2.3)
Dementia	5491 (5.3)
Major depression	22 378 (21.5)
Substance abuse (non-alcohol)	4164 (4.0)
Median number of medications (IQR)	8 (6–12)
Hospitalized at least once	26 921 (25.9)
Anticoagulation control	
Per cent time in range (TTR), mean (SD)	61.6 (22.1)

IQR, interquartile range. *Patients whose main indication for anticoagulation was valvular heart disease or prosthetic heart valve were excluded from this study.

Dependent variable: site-level anticoagulation control

Our dependent variable was mean site risk-adjusted TTR (also known as observed minus expected, or O-E). First, for each patient, we calculated the observed TTR (O) and applied the risk adjustment model to calculate the expected TTR (E). Then, an observed minus expected (O-E) score was calculated for each patient. The mean O-E score for each site constituted its risk-adjusted TTR. We also calculated the mean O, E and O-E for each site of care.

Independent variable: site mean INR value

As explained above, we included all INR values that occurred when a patient was 'on warfarin' according to our definition, and that were within 56 days of another INR (a 'valid interval'). We limited this study to patients who had at least four valid intervals for calculating per cent time in therapeutic range (TTR) [21]. For this purpose, a valid interval consists of two INR values separated by 56 days or less, without an intervening hospitalization. In this way, we ensured that patients had sufficient INR values to reliably indicate their target INR. For each patient, we calculated his or her mean INR value over the 2-year study period. We also calculated separate mean INR values in years 1 and 2 of the study for patients who had at least four valid intervals in one year. We computed, for each site, the overall mean INR value, as well as the proportions of patients with a mean INR value < 2.3, 2.3–2.7 and > 2.7.

INR values may be measured several times in rapid succession when the INR is low or high. We were concerned that this phenomenon might impact our assessment of patient- or site-level mean INR values. We therefore computed an alternative version of our independent variable, omitting all INR values obtained within 7 days after the previous value. The results of our analyses were essentially unchanged.

Statistical analyses

We examined the baseline characteristics of patients in our database. We calculated unadjusted and risk-adjusted TTR for each patient and for each site of care. We characterized each patient's mean INR value, each site's mean INR value, and the proportion of patients at each site with a mean INR of 2.3–2.7. We examined the ability of patient-level mean INR value to predict unadjusted and adjusted TTR using ANOVA, linear regression and cubic smoothing splines [23]. We examined the relationship between site-level proportion of patients with mean INR 2.3–2.7 and site-level anticoagulation control using linear regression. All analyses were conducted using SAS, version 9.1 (SAS Corporation, Cary, NC, USA).

Results

Patients

Our database contained 103 897 patients who were experienced users of warfarin for indications ordinarily requiring a target INR range of 2–3 (i.e. not mechanical heart valves). Patient characteristics are shown in Table 1. The sample was mostly male (98%) and had a median age of 72 years (IQR 62–79). More than half of the patients were anticoagulated for AF (64%), with many of the others anticoagulated for VTE (27%). Patients had a substantial burden of chronic disease: 40% had diabetes, 32% had heart failure, 29% had chronic lung disease, and 14% had chronic kidney disease. Patients also had a substantial burden of mental health disorders: for example, 22% had major depression, and 9% carried a diagnosis of alcohol abuse. As might be expected with this burden of co-morbidity, patients received many medications (a median of eight) and 26% were hospitalized at least once during the study.

Overall anticoagulation control was fair for this population of experienced users of warfarin (mean TTR = 61.6%).

Mean INR values – patient level

Patients varied widely in their mean INR values. The mean of all INR values in the database was 2.42, but 10% of patients were below 2.00 and 10% were above 2.83, so extreme deviations were not uncommon. The maximum predicted TTR occurred at a mean INR value of 2.43. At this value, patients had a predicted TTR of 71%; for each deviation of 0.1 from this value, the predicted TTR was 3.9% lower ($P < 0.001$). However, this relationship was not symmetric, with greater decrements in TTR occurring below a mean INR of 2.43 than above it (Table 2). An ANOVA test, comparing deciles of patients with regard to mean INR values, demonstrated a similar phenomenon (Table 2). Anticoagulation control was generally excellent among patients with a mean INR between 2.3 and 2.7, but quite poor outside of this range.

Mean INR values – site level

There were 100 sites of care in the database. The number of patients per site ranged from 74 to 4371 (mean 1039 per site). Site TTR varied from a low of 43% to a high of 72%, and site risk-adjusted TTR varied from 18% below to 12% above predicted. Site mean INR varied widely, from 2.22 to 2.89. The median site had 34% of patients with a mean INR < 2.3, 48% of patients with a mean INR 2.3–2.7, and 16% of patients with a mean INR > 2.7. However, there was considerable variation in this regard; sites ranged from 30 to 64% of patients with a mean INR of 2.3–2.7 (Fig. 1). In a regression analysis, for every 10% of patients at a site with a mean INR of 2.3–2.7, site-unadjusted TTR was 4.4% higher, and site risk-adjusted TTR was 3.8% higher ($P < 0.001$ for both findings).

The proportion of patients at each site with a mean INR 2.3–2.7 was consistent between years: comparing FY07 with FY08, the coefficient of correlation was 0.78 ($P < 0.001$). Sites were also somewhat consistent regarding whether their patients achieved mean INR values above or below this desirable range. For example, regarding the proportion of patients with a mean INR below 2.3, sites varied from 19 to 62%. The coefficient of correlation between FY07 and FY08 regarding the proportion with mean INR below 2.3 was 0.45 ($P < 0.001$). Several sites had a consistent pattern of having many patients with low mean INR. For example, sites KM and GD had among the very highest proportion of patients with a low mean INR in both years (KM, 56% and 65%; GD, 55% and 60%). The correlation between years for the proportion with a high mean INR was also appreciable ($r = 0.35$, $P < 0.001$), but not as strong a relationship.

Discussion

We examined mean INR values at the patient level and the site of care level. We had three main findings: (i) site mean INR varied widely; (ii) proportion of patients near 2.5 varied widely by site, and proportions low, high and near 2.5 were stable across years; and (iii) a higher proportion near 2.5 was strongly associated with higher site-level TTR. Several related points remain speculation: (i) intentional pursuit of non-standard INR target ranges is the principal cause of variation in site-level mean INR; (ii) actively discouraging clinicians from adopting non-standard targets will increase TTR and reduce adverse events; (iii) it will be possible to induce clinicians to change their practise regarding INR target ranges; and (iv) discouraging non-standard targets will not harm certain classes of patients, such as the oldest old (age 80+).

We will soon launch an initiative to improve mean TTR in the VA from its current 58% to 70% [9], an improvement that could prevent thousands of adverse events [24]. In light of the present study, one strategy will be to discourage sites from pursuing non-standard target INR ranges. We hypothesize that decreased use of non-standard target ranges will increase the proportion of patients with a mean INR of 2.3–2.7, and that sites that change this parameter the most will increase TTR the most. In addition, by directly measuring outcomes among important subgroups of patients (e.g. the oldest old), we hope to provide reassurance that avoiding non-standard target ranges benefits them as well, or at least does not harm them. This forthcoming study will therefore address whether the proportion of patients whose mean INR is near 2.5 can serve as a useful process measure in anticoagulation care.

The finding that extremes of mean INR are associated with poor anticoagulation control at the patient level is not surprising. However, the finding that sites of care vary more

Table 2 Patient-level analysis of relationship between mean INR value and per cent time in therapeutic range (TTR). Patients are divided into deciles by mean International Normalized Ratio (INR) values. The range of mean INRs is given for each decile (n = 103 897)

Mean INR group	Lowest	2	3	4	5	6	7	8	9	Highest
Range of mean INR	< 2.00	2.00–2.16	2.16–2.25	2.25–2.33	2.33–2.40	2.40–2.48	2.48–2.56	2.56–2.66	2.66–2.83	> 2.83
Observed TTR	30.6%	57.4%	65.9%	68.8% A	70.6% B	70.9% B	70.8% B	69.1% A	64.7%	46.8%
Adjusted TTR	−28.4%	−3.7%	+4.1%	+6.6% C	+8.2% D	+8.5% D	+8.4% D	+6.9% C	+2.9%	−13.4%

Observed TTR is between 0 and 100%. Adjusted TTR measures the difference between observed and expected TTR. Expected TTR is calculated for each patient based on his or her demographics and co-morbid conditions. Positive values for adjusted TTR indicate that a patient's observed TTR exceeded the expected TTR. ANOVA tests are significant for comparisons of observed TTR and adjusted TTR ($P < 0.001$ for both). Groups designated by a common letter are not significantly different from one another by Tukey's *post hoc* test.

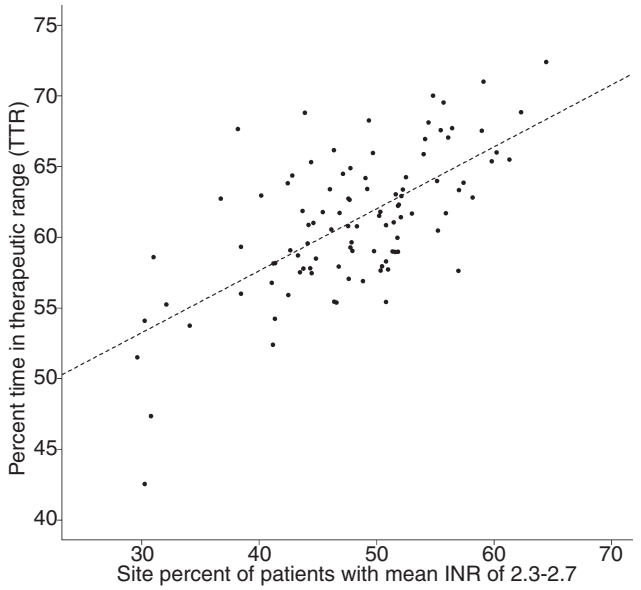


Fig. 1. Site-level per cent time in therapeutic range (TTR) plotted against site-level proportion of patients with mean INR of 2.3–2.7. (n = 100 sites).

than 3-fold regarding the proportion of patients with a mean INR below 2.3 suggests real differences in site-level practise. Our findings suggest that some sites routinely choose a guideline-concordant target range of 2–3 while others are likely to attempt to 'shade' patients below this range. This might take the form of an explicit designation of a non-standard target range (for example, 2–2.5) or an implicit tendency to aim for the lower end of a nominal target range of 2–3.

As discussed above, considerable work is needed to extend the results of this study and to clarify their implications. For now, however, we recommend that clinicians aim for standard target ranges, as are already recommended by prominent guidelines such as ACCP [5]. We are not advocating a target range of 2.3–2.7, or any other narrow range [19]. Clinicians should be equally satisfied with an INR of 2.1 or 2.9, provided the patient is under stable control. Rather, we recommend eliminating target ranges other than 2–3 and 2.5–3.5 [5,17,18,25,26].

Our study has important strengths, particularly the size and clinical detail of the database. However, we also acknowledge limitations. First, as stated above, we cannot know for sure why sites achieved such disparate mean INR values. While we suspect that it is a reflection of different styles of practise at each site, our observational database study is not equipped to confirm this suspicion. Second, we examined the impact of mean INR values on anticoagulation control, but did not examine the impact of mean INR values on adverse events. However, multiple studies have already linked TTR to adverse events [1–4], demonstrating the importance of pursuing a higher TTR both for populations and for individuals. In addition, multiple studies have also linked the pursuit of 'low' target INR ranges to inferior patient outcomes [11–15], as discussed above. Nevertheless, this study could have provided even more compelling evidence by directly linking site-level target ranges to site-level rates of adverse events. Unfortunately, we lack the data to directly examine adverse events with this dataset. In the future, we hope to examine this link more directly. Third, VA patients are mostly male and have higher rates of co-morbid illness, mental illness, substance abuse disorders and socioeconomic disadvantages than the general US population. This may have contributed to INR target range choices; for example, clinicians may justifiably fear hemorrhage in an alcoholic patient and therefore target a low INR value. However, these issues are unlikely to impact the underlying relationships we demonstrated between site-level mean INR values and site-level TTR. In addition, our risk-adjustment model should have accounted for the impact of these patient characteristics upon TTR, at least within the VA system [9].

In conclusion, 100 sites of care in the nation's largest integrated healthcare system varied considerably in their mean INR values and in the proportion of patients with a mean INR of 2.3–2.7. Our results suggest that pursuing standard INR target ranges (2–3 for most patients) could improve anticoagulation control, as measured by TTR. In a forthcoming intervention study, we plan to test this proposition by actively promoting the use of standard INR targets and observing the effects on TTR and patient outcomes.

Disclaimer

The opinions expressed in this manuscript do not necessarily represent the official views of the Department of Veterans Affairs.

Acknowledgements

This project was supported by a grant from the Veterans Affairs Health Services Research and Development Service (IIR-10-374). A. J. Rose is supported by a Career Development Award from the Veterans Affairs Health Services Research and Development Service (CDA-2-08-017). The sponsor had no role in the design and conduct of the study, the collection, management, analysis and interpretation of the data, and the preparation, review and approval of the manuscript.

Disclosure of Conflict of Interests

E. M. Hylek has received honoraria from Bayer and Bristol Myers Squibb, and has served on advisory boards for Boehringer-Ingelheim, Bristol Myers Squibb, Merck and Sanofi Aventis. None of the other authors report any potential conflict of interests.

References

1 Connolly SJ, Pogue J, Eikelboom J, Flaker G, Commerford P, Franzosi MG, Healey JS, Yusuf S. Benefit of oral anticoagulant over

antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation* 2008; **118**: 2029–37.
2 van Leeuwen Y, Rosendaal FR, Cannegieter SC. Prediction of hemorrhagic and thrombotic events in patients with mechanical heart valve prostheses treated with oral anticoagulants. *J Thromb Haemost* 2008; **6**: 451–6.
3 Veeger NJ, Piersma-Wichers M, Tijssen JG, Hillege HL, van der Meer J. Individual time within target range in patients treated with vitamin K antagonists: main determinant of quality of anticoagulation and predictor of clinical outcome. A retrospective study of 2300 consecutive patients with venous thromboembolism. *Br J Haematol* 2005; **128**: 513–9.
4 White HD, Gruber M, Feyzi J, Kaatz S, Tse HF, Husted S, Albers GW. Comparison of outcomes among patients randomized to warfarin therapy according to anticoagulant control: results from SPORTIF III and V. *Arch Intern Med* 2007; **167**: 239–45.
5 Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 2008; **133**: 160S–98S.
6 Ansell JE, Oertel LB, Wittkowsky AK, editors. *Managing Oral Anticoagulation Therapy: Clinical and Operational Guidelines*, 3rd edn. St Louis, MO: WoltersKluwer Health, 2009.
7 Garcia DA, Witt DM, Hylek E, Wittkowsky AK, Nutescu EA, Jacobson A, Moll S, Merli GJ, Crowther M, Earl L, Becker RC, Oertel L, Jaffer A, Ansell JE. Delivery of optimized anticoagulant therapy: consensus statement from the anticoagulation forum. *Ann Pharmacother* 2008; **42**: 979–88.
8 Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. *N Engl J Med* 1996; **335**: 966–70.
9 Rose AJ, Hylek EM, Ozonoff A, Ash AS, Reisman JI, Berlowitz DR. Risk-adjusted percent time in therapeutic range as a quality indicator for outpatient oral anticoagulation: results of the veterans affairs study to improve anticoagulation (VARIA). *Circ Cardiovasc Qual Outcomes* 2011; **4**: 22–9.
10 Rose AJ, Hylek EM, Berlowitz DR, Ash AS, Reisman JI, Ozonoff A. Prompt repeat testing after out-of-range INR values: a quality indicator for anticoagulation care. *Circ Cardiovasc Qual Outcomes* 2011; **4**: 276–82.
11 Stroke Prevention in Atrial Fibrillation Investigators. Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation: stroke prevention in atrial fibrillation III randomised clinical trial. *Lancet* 1996; **348**: 633–8.
12 Gullov AL, Koefoed BG, Petersen P. Bleeding during warfarin and aspirin therapy in patients with atrial fibrillation: the AFASAK 2 study. Atrial fibrillation aspirin and anticoagulation. *Arch Intern Med* 1999; **159**: 1322–8.
13 Hellemons BS, Langenberg M, Lodder J, Vermeer F, Schouten HJ, Lemmens T, van Ree JW, Knottnerus JA. Primary prevention of arterial thromboembolism in non-rheumatic atrial fibrillation in primary care: randomised controlled trial comparing two intensities of coumarin with aspirin. *BMJ* 1999; **319**: 958–64.
14 Kearon C, Ginsberg JS, Kovacs MJ, Anderson DR, Wells P, Julian JA, MacKinnon B, Weitz JI, Crowther MA, Dolan S, Turpie AG, Geerts W, Solymoss S, van Nguyen P, Demers C, Kahn SR, Kassis J, Rodger M, Hambleton J, Gent M. Comparison of low-intensity warfarin therapy with conventional-intensity warfarin therapy for long-term prevention of recurrent venous thromboembolism. *N Engl J Med* 2003; **349**: 631–9.
15 Pengo V, Palareti G, Cucchini U, Molinatti M, Del Bono R, Baudo F, Ghirarduzzi A, Pegoraro C, Iliceto S. Low-intensity oral anticoagulant plus low-dose aspirin during the first six months versus standard-intensity oral anticoagulant therapy after mechanical heart valve replacement: a pilot study of low-intensity warfarin and aspirin in cardiac prostheses (LIWACAP). *Clin Appl Thromb Hemost* 2007; **13**: 241–8.
16 Hull R, Hirsh J, Jay R, Carter C, England C, Gent M, Turpie AG, McLoughlin D, Dodd P, Thomas M, Raskob G, Ockelford P. Different intensities of oral anticoagulant therapy in the treatment of proximal-vein thrombosis. *N Engl J Med* 1982; **307**: 1676–81.
17 Kearon C, Kahn SR, Agnelli G, Goldhaber S, Raskob GE, Comerota AJ. Antithrombotic therapy for venous thromboembolic disease: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 2008; **133**: 454S–545S.
18 Singer DE, Albers GW, Dalen JE, Fang MC, Go AS, Halperin JL, Lip GY, Manning WJ. Antithrombotic therapy in atrial fibrillation: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 2008; **133**: 546S–92S.
19 Meier DJ, Seva S, Fay WP. A comparison of anticoagulation results of patients managed with narrow vs. standard international normalized ratio target ranges. *J Thromb Haemost* 2007; **5**: 1332–4.
20 RAND Health Corporation. ACOVE-2 Quality Indicators. Available at http://www.rand.org/content/dam/rand/www/external/health/projects/acove/docs/acove_qimedadmin.pdf. Accessed 26 May 2011.
21 Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thromb Haemost* 1993; **69**: 236–9.
22 Veterans Health Administration. *VHA Directive 2010-020. Anticoagulation Therapy Management*. Available at www.va.gov/vhapublications/viewpublication.asp?pub_ID = 2234. Accessed 9 January 2012.
23 Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. New York, NY: Chapman and Hall, 1990.
24 Rose AJ, Berlowitz DR, Ash AS, Ozonoff A, Hylek EM, Goldhaber-Fiebert JD. The business case for quality improvement: oral anticoagulation for atrial fibrillation. *Circ Cardiovasc Qual Outcomes* 2011; **4**: 416–24.
25 Keeling D, Baglin T, Tait C, Watson H, Perry D, Baglin C, Kitchen S, Makris M. Guidelines on oral anticoagulation with warfarin – fourth edition. *Br J Haematol* 2011; **154**: 311–24.
26 Salem DN, O'Gara PT, Madias C, Pauker SG. Valvular and structural heart disease: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 2008; **133**: 593S–629S.

HSR Health Services Research

DOI: 10.1111/j.1475-6773.2011.01377.x
RESEARCH ARTICLE

Organizational Characteristics of High- and Low-Performing Anticoagulation Clinics in the Veterans Health Administration

Adam J. Rose, Beth Ann Petrakis, Patricia Callahan, Scott Mambourg, Dimple Patel, Elaine M. Hylek, and Barbara G. Bokhour

Objective. Anticoagulation clinics (ACCs) can improve anticoagulation control and prevent adverse events. However, ACCs vary widely in their performance on anticoagulation control. Our objective was to compare the organization and management of top-performing with that of bottom-performing ACCs.

Data Sources/Study Setting. Three high outlier and three low outlier ACCs in the Veterans Health Administration (VA).

Study Design. Site visits with qualitative data collection and analysis.

Data Collection/Extraction Methods. We conducted semi-structured interviews with ACC staff regarding work flow, staffing, organization, and quality assurance efforts. We also observed ACC operations and collected documents, such as the clinic protocol. We used grounded thematic analysis to examine site-level factors associated with high and low outlier status.

Principal Findings. High outlier sites were characterized by (1) adequate (pharmacist) staffing and effective use of (nonpharmacist) support personnel; (2) innovation to standardize clinical practice around evidence-based guidelines; (3) the presence of a quality champion for the ACC; (4) higher staff qualifications; (5) a climate of ongoing group learning; and (6) internal efforts to measure performance. Although high outliers had all of these features, no low outlier had more than two of them.

Conclusions. The top-performing ACCs in the VA system shared six relatively recognizable characteristics. Efforts to improve performance should focus on these domains.

Key Words. Qualitative research, quality of health care, anticoagulants, pharmacists, organization and administration

Warfarin (Coumadin) is an oral anticoagulant given to millions of patients to treat or prevent blood clots. Warfarin requires routine blood tests to ensure that patients are sufficiently anticoagulated to prevent blood clots, but not so

thoroughly anticoagulated as to place them at excessive risk of bleeding. Warfarin can be difficult to manage, and clinicians must carefully monitor and adjust doses to keep the patients in the target range. Management in a dedicated anticoagulation clinic (ACC) can improve anticoagulation control (van Walraven et al. 2006). Veterans Health Administration (VA) medical centers are required to manage all patients receiving warfarin in an ACC, which is generally run by clinical pharmacists (Veterans Health Administration 2009).

Nevertheless, even among patients receiving treatment in an ACC, there is substantial variation in anticoagulation control (Rose et al. 2011a). Many patients have poor control (van Walraven et al. 2006; Rose et al. 2010), which increases rates of stroke, venous thromboembolism, major hemorrhage, and death (Veeger et al. 2005; White et al. 2007; Wallentin et al. 2010). Anticoagulation control can be measured by the percentage of time in therapeutic range (TTR) (Rosendaal et al. 1993). We have proposed using TTR as a quality measure (Rose et al. 2009). In our recent study, mean site TTR varied among 100 VA ACCs by more than 20 percent, a large and clinically important difference (Rose et al. 2011a).

We sought to understand this variation in performance using the positive deviance approach (Bradley et al. 2009). This approach emphasizes in-depth qualitative study of organizations with exceptionally high performance to understand the factors that contribute to their excellence (Peters and Waterman 1982; Spear and Bowen 1999; Gawande 2007). Qualitative methods are ideal for studying organizational culture, which can be difficult to characterize using quantitative methods (Sofaer 1999; Patton 2002). The Anticoagulation Forum has proposed nine key domains (Table 1) that are essential to establishing and maintaining a high-quality ACC (Garcia et al. 2008). We sought to explore these nine domains, while remaining open to others that would emerge through this qualitative study.

Address correspondence to Adam J. Rose, M.D., M.Sc., F.A.C.P., Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, 200 Springs Road, Building 70, Bedford, MA 01730; e-mail: adamrose@bu.edu. Adam J. Rose, M.D., M.Sc., F.A.C.P., Beth Ann Petrakis, M.P.A., Elaine M. Hylek, M.D., M.P.H., and Barbara G. Bokhour, Ph.D., are with the Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, Bedford, MA. Adam J. Rose, M.D., M.Sc., F.A.C.P., and Elaine M. Hylek, M.D., M.P.H., are with the Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA. Adam J. Rose, M.D., M.Sc., F.A.C.P., Beth Ann Petrakis, M.P.A., Patricia Callahan, RP.h., Scott Mambourg, PharmD, B.C.P.S., Dimple Patel, PharmD, B.C.P.S., Elaine M. Hylek, M.D., M.P.H., and Barbara G. Bokhour, Ph.D., are with the United States Department of Veterans Affairs, Washington, DC. Barbara G. Bokhour, Ph.D., is with the Department of Health Policy and Management, Boston University School of Public Health, Boston, MA.

Table 1: Nine Recommendations for Optimal Anticoagulation Care as per the Anticoagulation Forum Consensus Statement (Garcia et al. 2008)

Domain	*Explanation*
1. Qualifications of personnel	ACC staff should be licensed health care professionals possessing core competency related to anticoagulation therapy.
2. Supervision	The ACC should have a well-defined relationship with responsible/referring health care providers (e.g., the primary care physician).
3. Care management and coordination	The ACC should have written policies and procedures that are approved by the clinic's medical director. An efficient system for scheduling and tracking patients should be utilized.
4. Documentation	The ACC should have an accurate system of documentation to ensure that clinically relevant data are available to staff at all times.
5. Patient education	The delivery of anticoagulation care should address the educational needs of patients and their caregivers.
6. Patient selection and assessment	Optimized anticoagulant therapy should be instituted only after careful consideration of the risk and benefit for an individual patient. The appropriateness of anticoagulation therapy should be periodically reviewed for each patient.
7. Laboratory monitoring	Optimized anticoagulation therapy should incorporate regular laboratory monitoring of anticoagulant effect, using either a high-quality laboratory or a well-maintained point of care device.
8. Initiation of therapy	The initiation of optimized anticoagulation therapy should use a systematic, evidence-based approach.
9. Maintenance of therapy	The delivery of optimized anticoagulation should use a systematic process for longitudinal patient assessment, adjustment of anticoagulant drug doses, scheduling of follow-up visits, and interruptions of therapy for elective procedures.

METHODS

Selection of Study Sites

Based on our rankings of risk-adjusted anticoagulation control at 100 VA sites (Rose et al. 2010, 2011a), we selected 3 of the top 10 and 3 of the bottom 10 sites, and obtained IRB approval from all study sites. Sites were not informed of their high or low outlier status. Although we selected sites based on their performance from October 2006 to September 2008 (the "measurement period"), site visits occurred between May 2010 and March 2011. We took care to identify personnel and organizational factors that had changed since the measurement period. To a large extent, these factors and personnel had not changed, but when they had, we viewed this as an additional opportunity to examine not only static models of high and low performance but also changes in performance. When conditions as we found them differed in important ways from what was present during the measurement period, we have noted this in the text.

Data Collection

Interviews. One author (A. J. R.) visited each study site for two working days. At each site, we sought to interview all staff members who work in the ACC or whose work impacts the ACC. All employees agreed except one, and all participants gave informed consent. Each participant was interviewed in private for 20–60 minutes using a semi-structured interview technique. Interview topics were based largely upon the Anticoagulation Forum recommendations (Garcia et al. 2008) and included ACC work flow, staffing, organization, and quality assurance efforts. We also asked about the clinic's patient population, the challenges of managing anticoagulation, and the participant's opinion of the clinic's performance. All interviews were audio-recorded and transcribed verbatim.

Observation. At each site, we observed approximately 4 hours of direct clinical care, which occurred face-to-face or by telephone. We observed delivery of care, work environment, work pace and flow, patient-provider interactions, and interactions among pharmacists and pharmacy technicians working in the ACC.

Documents. We obtained ACC-related documents, including ACC protocols, note and consult templates from the electronic medical record, training manuals for ACC staff, education materials for non-ACC staff, patient education materials, patient-provider treatment agreements, form letters sent to patients, quality assurance forms and reports, and screen shots of clinical management software used by the ACC.

Data Analysis

Using grounded thematic analysis, we coded interview transcripts, identifying broad "domains" of ACC organization and subtopics within each domain (Patton 2002). Having coded the interview transcripts, we added information gleaned from the field notes and the ACC documents. We synthesized the data from each site and produced a profile of each site's organization and management. As successive study sites were added, we used constant comparison to contrast our findings from various sites, noting similarities and differences between high- and low-performing sites. We made a conscious effort to note counterexamples, particularly positive aspects of bottom-performing sites and negative aspects of top-performing sites. We identified key features present or

absent at each site, noting relationships between these features and site performance. Finally, we performed internal member checking among our co-authors, many of whom have direct experience managing oral anticoagulation, to ensure that our findings resonate with anticoagulation practitioners.

RESULTS

The six sites we visited represented six different states and all geographic regions of the United States. Details about site performance are located in Table 2. Table 3 lists the job descriptions of the 55 interviewees.

We identified 11 domains related to site organization; six were related to site performance and five were unrelated (Table 4). Although high outliers had all six of the features related to performance, no low outlier had more than two of them. The important domains were (1) adequate (pharmacist) staffing levels and effective use of (nonpharmacist) support personnel; (2) innovations to standardize clinical practice around current evidence and guidelines; (3) the presence of a quality champion; (4) higher staff qualifications; (5) a climate of ongoing group learning; and (6) internal performance measurement.

Domain #1: Staffing

We observed, and were told by interviewees, that ACC staff at the low outlier sites had a rushed work pace and a chaotic work environment, whereas

Table 2: Three High Outlier and Three Low Outlier Sites, as Measured by Performance on Percent Time in Range (TTR), a Measure of Anticoagulation Control

Site Number	*Outlier Status*	*TTR, 2007–2008 (%)*
1	High	67
2	High	67
3	High	70
4	Low	52
5	Low	54
6	Low	54

Notes. Although Sites Were Selected for Inclusion Based on Risk-Adjusted TTR (Rose et al. 2011b), Unadjusted TTR Is Presented Here for Ease of Understanding. TTR was computed using the Rosendaal method (Rosendaal et al. 1993). TTR above 60 percent is generally considered adequate, whereas TTR above 70 percent is generally considered excellent.

Table 3: Type of Staff Interviewed at Study Sites

	Interviews, n
Direct-care ACC staff	
Pharmacist/ACC coordinator	5
Pharmacist, not ACC coordinator	19
Pharmacy resident	1
ACC support staff	
Pharmacy technician	3
Clerk/secretary	2
Nurse	1
Pharmacy administration	
Chief of pharmacy	5
Associate chief of pharmacy	4
Middle manager, pharmacy	2
Physicians (ACC medical director)	
Staff physician	1
Chief of primary care	3
Chief of medicine	1
Chief of cardiology	1
Other staff	
Laboratory supervisor	4
Medical director of laboratory	1
Phlebotomist	1
Nurse practitioner, cardiology	1
Supervisor of clerks	1
Total	55

staff at the high outliers worked at a comfortable pace in a well-organized environment. Based on expert opinion, consensus guidelines for organizing an ACC recommend a ratio of no more than 400 patients per pharmacist full-time equivalent (FTE) (Ansell et al. 2008). Details about staffing at our six study sites are in Table 5. In general, the adequacy of staffing was not fully characterized by the patient/FTE ratio but also depended on the effective integration of support staff into the ACC, the degree to which staff could focus on their ACC duties without other responsibilities, and avoiding an inappropriate reliance upon trainees (pharmacy residents) to provide manpower, as will be discussed below.

Pharmacists. Anticoagulation clinic staff at low outlier sites consistently reported being rushed due to inadequate staff allocations; we also observed this rushed work pace. In contrast, staff at high outliers only rarely reported mild issues with staffing related to vacations and sick leave. Understaffing at

Table 4: Summary of Six Anticoagulation Clinics on Six Domains Related to Site Performance and Five Domains Not Related to Site Performance

	High Outliers			*Low Outliers*		
	Site 1	*Site 2*	*Site 3*	*Site 4*	*Site 5*	*Site 6*
Domains related to performance						
Staffing	✓	✓	✓			
Innovation	✓	✓	✓		✓	
Champion	✓	✓	✓			
Qualifications	✓	✓	✓			✓
Group learning	✓	✓	✓			✓
Measurement	✓	✓	✓			
Domains not related to performance						
Telephone clinic		✓	✓			✓
Commitment	✓	✓	✓	✓	✓	✓
Integration	✓	✓	✓	✓	✓	✓
Patient selection						
Tenure		✓	✓	✓	✓	✓

Key to table:
✓ = present during measurement period of 2007–2008.
No symbol = absent during measurement period of 2007–2008.
Key to domains:
Staffing: Sufficient staffing to handle workload without rushing.
Innovation: Innovations to encourage evidence-based practice.
Champion: Presence of one or more quality champions for the ACC.
Qualifications: Hiring residency-trained pharmacists.
Group learning: Creating a climate of ongoing group learning.
Measurement: Internal performance measurement.
Telephone clinic: Majority of patients managed via telephone rather than face-to-face.
Commitment: Demonstrated commitment to serving veterans.
Integration: Effectively leverages the advantages of membership in an integrated health system.
Patient selection: Selectively refusing to treat difficult or troublesome patients (gaming).
Tenure: Clinic established more than 20 years ago.

low outliers was manifested through frequently overbooked appointments, chaotic work days, and sagging morale. An interview with one pharmacist revealed the stress prevalent at the low outlier sites:

> Q: Are you pretty happy with how your job here is going?
> A: I'm overwhelmed…over the years, our numbers have grown … and no one has really looked at our workload and I think we're really stretched. (Pharmacist, Low Outlier)

A pharmacist at another low outlier suggested that being short-staffed may prevent them from following some patients as frequently as needed, a potentially serious issue with quality of care:

Table 5: Details for Clinic Staffing in Six VA ACCs

Panel Size		*FTE*	*Patient/ FTE Ratio*	*Support Staff*	*Details*
High outlier sites					
Site 1	550	1.5	375	Specially trained clerks	ACC providers have considerable time for administrative tasks, much of which is also devoted to the ACC
Site 2	1700	4	425	2 pharmacy technicians	Workflow is highly streamlined and is aided by clinical management software
Site 3	250	1.5	167	None	
Low outlier sites					
Site 4	600	2	300	None	Pharmacists not dedicated to ACC; distraction may reduce effectiveness
Site 5	1000	1	1000	None	Approximately 1 additional FTE of resident labor in ACC
Site 6	800	1.5	533	1 pharmacy technician	

Notes. Panel Sizes Were Determined from VA Automated Data (Rose et al. 2011b), whereas Pharmacist FTEs Assigned to the ACC during the Measurement Period (10/06-9/08) Were Determined through Interviews and Direct Observation. A Patient to FTE Ratio of No More Than 400 Is Generally Recommended by Consensus Guidelines (Ansell et al. 2008).
ACC, anticoagulation clinic; FTE, full-time equivalents (pharmacists, not counting support staff); VA, Veterans Health Administration.

> Sometimes we need closer follow-up, but we don't do closer follow-up 'cause we don't have enough help. So I think we're doing the best we can with what we've got. (Pharmacist, Low Outlier)

By contrast, ACC providers at the top sites had fewer patients per FTE, which was reflected in their work pace. One participant noted:

> I don't think we're capacity most days. I don't think we have 40 face-to-face visits most days, but that's good, because that allows us to see patients more frequently if necessary. There are still appointments left to see those patients without overbooking. (Associate Chief of Pharmacy, High Outlier)

Support Staff. Adequate staffing also relates to the role of support staff, individuals who work in the ACC but do not have the authority to prescribe warfarin. When present, pharmacy technicians were universally lauded as contributing to the smooth functioning of the ACC. For example, at one high outlier, four FTE of pharmacists were supplemented by two FTE of pharmacy technicians, who handle incoming calls, distribute workload, and mail letters,

thereby freeing the pharmacists to concentrate on higher level tasks, such as managing patients who are out of range. However, personnel other than pharmacy technicians can also provide effective support. At another high outlier, only the most experienced clerks are given special training and permission to schedule appointments for the ACC. Low outliers were generally characterized by ineffective approaches to support staffing: either untrained clerks or, at two sites, no support staff at all.

Dedicated Time for ACC Duties. At most study sites, ACC pharmacists had dedicated days when they staffed the ACC without other obligations. In contrast, at one low outlier, ACC pharmacists were simultaneously responsible for anticoagulation and other tasks related to primary care pharmacy. Pharmacists at this site were frequently interrupted from ACC duties to attend to non-ACC matters.

For example, we observed a pharmacist from this low outlier site treating a patient who had suffered a blood clot in the lung. The patient was hard of hearing, possibly suffered from dementia, lived in a trailer, and did not fully understand or accept the need for anticoagulation. The visit was interrupted six times by people walking into the room (often without knocking) to ask the pharmacist questions in her role as a primary care team member. Each interruption caused a loss of focus and momentum, and the overall effect was to prolong the visit and to transform an already difficult encounter into an almost insurmountable challenge.

Avoiding Inappropriate Use of Pharmacy Residents. At most of the sites we visited, pharmacy residents were in the ACC primarily to learn and were not expected to contribute greatly to staffing due to their inexperience and need to be supervised. However, at one low outlier site, residents were expected to be about as productive as the senior pharmacists. They handled about half of the workload of the ACC and we were explicitly told that they constituted an important part of the staffing plan. A pharmacy resident from this site, who had recently rotated through the ACC for a month, remarked:

> I feel like the pharmacy residents do a lot here … just because there are so many patients … I'm interviewing patients in 10 or 15 minute increments, very fast. (Pharmacy Resident, Low Outlier)

The fact that half the care at this site was delivered by pharmacy residents may have contributed to uneven quality of clinical management. A pharmacist at this site admitted that the quality of the pharmacy residents can vary from day to day and from month to month: "It's hit and miss with them." This does not suggest that these residents or their training are lacking, merely that they need to complete residency before so much can be expected of them.

Domain #2: Innovations to Encourage Uniform, Evidence-Based Practice

Top sites developed innovative ways for encouraging uniformity of practice. In general, the organizing principles for standardization were drawn from existing clinical guidelines about anticoagulation care, especially the American College of Chest Physicians (ACCP) guidelines for managing warfarin (Ansell et al. 2008). This domain, with its emphasis on the systematic pursuit of uniform, evidence-based practice, is also embodied in the Anticoagulation Forum Consensus Guidelines for organizing an ACC (Garcia et al. 2008). We noticed at least three examples of how sites pursued uniform, evidence-based practice: developing note templates as an aid to clinical reasoning and documentation, adopting new software packages to enhance workflow, and developing systems to reduce loss to follow-up.

Evidence-Based Note Templates. One way to reinforce evidence-based practice is to design note templates around it. Although all of the sites had note templates, some were especially effective at encouraging adherence to clinical guidelines. For example, one site's ACC Coordinator created a note to guide clinicians through the decision-making process of how to manage a temporary interruption of warfarin therapy for a procedure. Although this was a low outlier site, this note was introduced after the measurement period as part of an effort to improve the ACC.

> Previously … if we heard about a patient … stopping [warfarin] for a procedure we would kinda wait and see if somebody had the idea of using bridge. I standardized that process … [now] if we hear a patient's stopping Coumadin he's gonna do it through a standardized review process that's based strictly on the [ACCP] guidelines. I created a note template that quotes what the guidelines say … and then we walk 'em through the peri-procedure risk assessment note which gets documented in the medical record and is sent to the doctor. (Pharmacist, Low Outlier)

Adopting New Software to Aid Clinical Practice. One high outlier site used a software package to help streamline anticoagulation management. Participants commented, and we observed, that the software contributed to this site's ability to handle a large patient load without compromising quality. The software improved work flow, helped make sure patients were not lost to follow-up, and allowed for internal performance measurement with a minimum of effort. A low outlier site also started using this software package in 2010 (after the measurement period) with similar positive results.

Reducing Loss to Follow-Up. It is important not to lose ACC patients to follow-up, because without proper monitoring, patients may be at risk for serious adverse events. All six sites had various techniques to avoid losing patients—the most common was to limit warfarin prescriptions to a short period of time (often 30 days), to ensure that the patient must follow-up. However, some sites went farther. At one high outlier, the ACC Coordinator "runs the list" monthly to manually search for patients who have not been seen for over 30 days (a laborious process). The two sites that had adopted the new software package easily performed this task, because the software monitors this automatically. One low outlier had no system to prevent loss to follow-up:

> Actually we don't have a system in place for that. Now, if we could run our clinics quarterly or have some kind of computer program where if a person hasn't been to see us in say, two months, then we could follow it, but right now we don't have a way of tracking 'em down. (Pharmacist, Low Outlier)

Note that this pharmacist recognizes the potential for manually running the list or a "computer program," elements that are present at the high outlier sites.

Domain #3: Quality Champion

The best sites had one or more strong quality champions for the ACC, although this champion could occupy many different formal positions. At one high outlier, this champion was the Associate Chief of Pharmacy, also the ACC's founder. She wrote all the procedures for the ACC, trained her replacements, and continued to take a special interest in the ACC's performance. At another high outlier, this champion was the ACC Coordinator,

while at a third, the main quality champion for the ACC was the Chief of Pharmacy. In the following quote, a quality champion explains her role with regard to the ACC:

> I would say my main thing is to make sure that I really kinda do the quality aspect…make sure that the…care that we're providing is appropriate. That if we're gonna offer a program, that it's doing what it needs to do, and if it isn't then what do I need to do to fix it. (Chief of Pharmacy, High Outlier)

None of the low outliers had a quality champion. For example, one low outlier had a middle manager in the pharmacy department whose job was to oversee credentialing, peer reviews, academic pursuits, and quality assurance for pharmacists. However, he had almost no role with regard to the ACC pharmacists, because they were located in primary care rather than pharmacy. Instead, these pharmacists were supervised by the director of primary care, a physician. Although she has made her primary care clinic into a model of performance, she saw the pharmacists primarily in their role as supporting cast for the primary care clinic and was relatively unaware of the pharmacists as the central movers in managing anticoagulation. Thus, there were two people qualified to serve as the quality champion for the ACC (the pharmacy middle manager and the ACC medical director), but one was not given the authority to do so and the other did not see it as part of her job.

Domain #4: Staff Qualifications

Many pharmacists complete a 1-year residency after obtaining their PharmD degree. The purpose of residency is to gain a set of skills in a mentored setting, especially the ability to interact with patients to directly manage chronic conditions. All of the ACC pharmacists at the top outlier sites had completed residencies in pharmacy. In contrast, none of the pharmacists at two of the low outlier sites were residency trained.

We observed important differences between pharmacists with and without residency training. The residency-trained pharmacists had greater facility communicating with patients. The pharmacists that lacked residency training struggled to handle difficult situations, often despite years of experience. At one site, we did observe a nonresidency trained pharmacist who was instead completing a protracted apprenticeship in the ACC. This suggests that other kinds of training can also help pharmacists to achieve a similar degree of facility with the skills needed for direct patient management.

Domain #5: Climate of Group Learning

Clinicians frequently discuss difficult cases to solicit the opinions and insights of their colleagues. At the high outlier sites, such interactions were explicitly encouraged, and they occurred frequently:

> So it's good to have other people next to you in the room you can bounce ideas off of. (Pharmacist, High Outlier)

These interactions were largely absent from the low outlier sites; providers at these sites tended to practice without the benefit of their colleagues' opinions. When asked about how often she seeks advice from colleagues, this pharmacist said:

> Not very often. I would probably have to say less than once every 3 months. I would say we don't tend to collaborate a lot. (Pharmacist, Low Outlier)

We directly observed a marked contrast between low and high outliers with regard to the frequency and quality of discussions of difficult cases. The availability of colleagues for discussion at the high outlier sites seemed to contribute not only to the management of the case at hand but also to an atmosphere of learning from each other.

Domain #6: Internal Performance Measurement

The top-performing sites emphasized internal performance measurement. One high outlier manually collected data at every patient encounter (a laborious task) to be able to calculate and track percentage of lab values in range and rates of adverse events in their ACC. The Associate Chief of Pharmacy, when asked about the strengths of his site's ACC, demonstrated detailed awareness of its performance:

> About two-thirds of our patients on average are within the therapeutic range. Our no-show and cancelation rate typically combined is less than 12 to 15 percent … thromboembolism is usually less than two percent, and major bleeding is usually less than 5 percent. (Associate Chief of Pharmacy, High Outlier)

At other sites, anticoagulation software or regionally produced performance reports made performance measurement particularly easy to accomplish. One high outlier site began using such software well before our

measurement period. A low outlier site recently adopted the same software. The ACC coordinator at this site commented:

> When we started using software as of January 2010, we kinda changed our QA process. We were now able to calculate our TTR or actually get reports from that from the software … we look at … thrombotic events, bleeding events … patients lost to follow-up … When the software took over we started doing that more routinely … it's something that has become a part of our practice. (Pharmacist, Low Outlier)

Low outlier sites did not measure performance in 2007–2008, although two of them have begun doing so since then. At the third low outlier, the ACC medical director admitted that they do not yet measure ACC performance in a rigorous, proactive way. When asked about ACC performance, she appeared to be unprepared for the question and thus uncomfortable with the topic. She focused on one-time episodes such as complaints or adverse events as a way to measure quality:

> You mean in terms of bad outcomes with the patients or in terms of? [Clears throat] Yes we definite—that, I would know [Clears throat]. We do have adverse outcomes. It's something that we look at. So if … a patient comes in with [a critical degree of overanticoagulation] and nobody did anything about it, I would know about this … if there's an adverse occurrence, then I'm told about this. (ACC Medical Director, Low Outlier)

In the absence of quality measurement, ACC pharmacists at this site believed that their clinic was doing very well—but in fact, their TTR was among the lowest of any VA site. One pharmacist guessed that her patients spend 70 percent of time in range, whereas another guessed 75 percent. In fact, TTR at this site was closer to 55 percent, a large and clinically important difference.

Five Domains Not Related to ACC Performance

There were at least five domains that were not associated with ACC performance, although we had expected that they might be (Table 4).

1. We had expected to find that quality might vary due to the configuration of the ACC, that is, telephone versus face-to-face management of patients and the use of point-of-care (fingerstick) testing versus reliance upon venous blood samples. We observed highly variable

configurations of care among our six sites. For example, three sites had predominantly telephone clinics and three saw their patients face-to-face, but neither model was associated with high or low outlier status.

2. All sites, whether low or high outliers, exhibited a strong commitment to serving veteran patients. Staff went to great lengths to help patients, contacting them after hours when necessary and even assisting them with matters unrelated to anticoagulation.
3. The VA's integrated system of care, and particularly the electronic medical record, are thought to contribute to the VA's generally high performance (Longman 2007). However, we did not notice any differences among sites regarding the extent to which they utilized these advantages.
4. We had suspected that high outlier sites might achieve better performance in part by being more selective in the patients they agree to manage (gaming). However, all sites went to extraordinary lengths to manage even the most challenging patients.
5. We had expected that ACCs established long ago might perform better than those established recently; in fact, only one of our ACCs was established less than 10 years before our site visit, and it was a high outlier.

In addition to these five domains, it should be noted that the nine domains identified as important by the Anticoagulation Forum (Garcia et al. 2008) generally were not sufficient to separate high- and low-performing sites (Table 1). Exceptions included Item 3, where high outlier sites generally had better systems for tracking patients and preventing loss to follow-up, and aspects of some of the other items, in that high outliers achieved greater uniformity of practice. In most respects, however, high and low outliers fulfilled the Anticoagulation Forum recommendations equally.

DISCUSSION

In this study, we visited three of the best-performing and three of the worst-performing ACCs in the VA system. We sought to understand how elements of ACC organization and management related to these differences in site-level performance. We found that six critical domains separated high and low outlier sites. Although we had anticipated some of these domains, our

investigation revealed details that we had not anticipated. For example, while we had suspected that staff credentials might be important, we had assumed that all ACC pharmacists would be residency trained, and that those at the top sites would be distinguished by additional credentials. We discovered the importance of residency training through its absence among staff at some low outlier sites and its apparent impact on their fluency as clinicians.

We also found that the recommendations of the Anticoagulation Forum (Garcia et al. 2008) were not sufficient to separate high outliers from our three low outliers. To a large extent, all six study sites had met all of these criteria. It is apparent, therefore, that the pursuit of excellence in oral anticoagulation care will require a comprehensive program of quality measurement and quality improvement (Rose et al. 2009).

In a previous study, we examined many of the same domains, including staffing, presence of quality improvement programs, and training of ACC staff but were unable to link them to site-level performance (Rose et al. 2011b). The previous study relied upon a site-level questionnaire to gather data, but complex ideas like adequate staffing may not easily be encapsulated in a single number. In the present study, we examined not only the FTE ratio but also the context and configuration of care delivery, the workflow, and the importance of support staff in allowing pharmacists to focus on higher level tasks. The present study suggests that adequate staffing is indeed important but that a simple numerical ratio of pharmacists to patients may not fully capture all the data needed to understand this issue.

Anticoagulation clinics improve anticoagulation control and usually reduce adverse events compared with "usual care" (van Walraven et al. 2006; Ansell et al. 2008), but less is known about how ACC management improves these outcomes. Our study suggests that an ACC should be characterized by a core group of well-qualified and well-trained staff, dedicated time to focus on anticoagulation (without other responsibilities), development of innovations to improve care, opportunities to discuss difficult cases and learn from each other, internal performance measurement, and a quality champion, whose role it is to facilitate all of these other things. To the extent that our low outlier sites lacked these features, they did not fully embody the ideal of an ACC. It should not be surprising, therefore, to find that anticoagulation control at these low outliers was similar to what has been reported for non-ACC settings (van Walraven et al. 2006).

Our study also suggests that building time into schedules for the purpose of problem solving and knowledge sharing can increase the value created per

person. Discussion of difficult cases allows for group learning and may even encourage group solutions to frequently encountered problems. In contrast, while providers at low outlier sites may learn from challenging cases as individuals, such learning is unlikely to spur practice innovations. It may not be necessary to add additional people to set aside time for these activities; rather, by emphasizing these activities, an ACC can foster a climate of efficiency in which people simply accomplish more and do it better.

Insights gained from this study are not limited to ACCs alone. In fact, Curry et al. found remarkably similar themes that differentiated between hospitals ranking in the top and bottom 5 percent on risk-adjusted mortality after acute myocardial infarction (Curry et al. 2011). Thus, there may be considerable overlap between the ingredients of success in a large program (such as treatment of myocardial infarction) and a small program (such as management of anticoagulation).

This study has important strengths. Our in-depth observation allowed us to uncover previously unsuspected findings, as discussed above. However, some limitations should also be noted. First, we were unable to gain access to some of the worst-performing ACCs in the VA system, primarily because those sites did not have an IRB of record, and therefore could not participate in research. We might have gained further insight had we been able to visit low outlier sites with worse performance. Second, a considerable amount of time elapsed between our measurement period and our site visits. This was unavoidable because of the time needed to identify outlier sites and obtain IRB approvals. Third, due to time and resource constraints, we were only able to visit six sites. We might have learned more if we had visited additional sites. Fourth, we realize that local circumstances may contribute to some of the suboptimal features we observed, such as a heavy reliance on pharmacy residents, high patient-FTE ratios, or a lack of residency-trained pharmacists. Strategies for addressing local resource limitations are beyond the scope of our study. Finally, the site visitor was not blinded to outlier status, and the analysis of data was also not blinded. However, we consciously tried to note what was suboptimal about high outlier sites and what was good about low outlier sites.

In summary, we found that six domains of organization and management are related to ACC performance. Clinics have the potential to transform themselves, and their performance, by focusing on changing these six domains for the better. This study reminds us that excellence is within our reach if we focus on the most important determinants of performance.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This study was funded by the VA Stroke Quality Enhancement Research Initiative (QUERI), Project RRP-10-054. Dr. Rose is supported by a VA HSR&D Career Development Award (CDA-2-08-017). The authors thank Russell Buono, Mitchell Finkel, Elizabeth Goy, and Anthony Vo for their help collecting the data. The authors thank Steve Spear, Rani Elwy, Dan Berlowitz, Arlene Ash, and Vicky Parker for their helpful comments on the manuscript. The authors thank all the interviewees for participating.

Disclosures: Dr. Hylek reports receiving honoraria from Bayer and Bristol Myers Squibb, as well as serving on advisory boards for Boehringer-Ingelheim, Bristol Myers Squibb, Merck, and Sanofi Aventis. None of the other authors report any potential conflicts of interest.

Disclaimers: The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

REFERENCES

Ansell, J., J. Hirsh, E. Hylek, A. Jacobson, M. Crowther, and G. Palareti. 2008. "Pharmacology and Management of the Vitamin K Antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition)." *Chest* 133 (6 suppl): 160S–98S.

Bradley, E. H., L. A. Curry, S. Ramanadhan, L. Rowe, I. M. Nembhard, and H. M. Krumholz. 2009. "Research in Action: Using Positive Deviance to Improve Quality of Health Care." *Implementation Science* 4: 25.

Curry, L. A., E. Spatz, E. Cherlin, J. W. Thompson, D. Berg, H. H. Ting, C. Decker, H. M. Krumholz, and E. H. Bradley. 2011. "What Distinguishes Top-Performing Hospitals in Acute Myocardial Infarction Mortality Rates? A Qualitative Study." *Annals of Internal Medicine* 154 (6): 384–90.

Garcia, D. A., D. M. Witt, E. Hylek, A. K. Wittkowsky, E. A. Nutescu, A. Jacobson, S. Moll, G. J. Merli, M. Crowther, L. Earl, R. C. Becker, L. Oertel, A. Jaffer, and J. E. Ansell. 2008. "Delivery of Optimized Anticoagulant Therapy: Consensus Statement from the Anticoagulation Forum." *Annals of Pharmacotherapy* 42 (July/August): 979–88.

Gawande, A. 2007. "The Bell Curve." In *Better*, pp. 201–30. New York: Metropolitan Books.

Longman, P. 2007. *Best Care Anywhere: Why VA Health Care Is Better Than Yours*. Sausalito, CA: PoliPointPress.

Patton, M. Q. 2002. *Qualitative Research and Evaluation Methods.* Thousand Oaks, CA: Sage Publications.

Peters, T. J., and R. H. Waterman. 1982. *In Search of Excellence: Lessons from America's Best-Run Companies.* New York, NY: HarperCollins.

Rose, A. J., D. R. Berlowitz, S. M. Frayne, and E. M. Hylek. 2009. "Measuring Quality of Oral Anticoagulation Care: Extending Quality Measurement to a New Field." *Joint Commission Journal on Quality and Patient Safety* 35 (3): 146–55.

Rose, A. J., E. M. Hylek, A. Ozonoff, A. S. Ash, J. I. Reisman, and D. R. Berlowitz. 2010. "Patient Characteristics Associated with Oral Anticoagulation Control: Results of the Veterans AffaiRs Study to Improve Anticoagulation (VARIA)." *Journal of Thrombosis and Haemostasis* 8 (10): 2182–91.

———. 2011a. "Risk-Adjusted Percent Time in Therapeutic Range as a Quality Indicator for Outpatient Oral Anticoagulation: Results of the Veterans Affairs Study to Improve Anticoagulation (VARIA)." *Circulation. Cardiovascular Quality and Outcomes* 4 (1): 22–9.

Rose, A. J., E. M. Hylek, A. Ozonoff, A. S. Ash, J. I. Reisman, P. P. Callahan, M. M. Gordon, and D. R. Berlowitz. 2011b. "Relevance of Current Guidelines for Organizing an Anticoagulation Clinic." *American Journal of Managed Care* 17 (4): 284–9.

Rosendaal, F. R., S. C. Cannegieter, F. J. van der Meer, and E. Briet. 1993. "A Method to Determine the Optimal Intensity of Oral Anticoagulant Therapy." *Thrombosis and Haemostasis* 69 (3): 236–9.

Sofaer, S. 1999. "Qualitative Methods: What Are They and Why Use Them?" *Health Services Research* 34 (5 Pt 2): 1101–18.

Spear, S., and H. K. Bowen. 1999. "Decoding the DNA of the Toyota Production System." *Harvard Business Review* September–October: 97–106.

Veeger, N. J., M. Piersma-Wichers, J. G. Tijssen, H. L. Hillege, and J. van der Meer. 2005. "Individual Time within Target Range in Patients Treated with Vitamin K Antagonists: Main Determinant of Quality of Anticoagulation and Predictor of Clinical Outcome. A Retrospective Study of 2300 Consecutive Patients with Venous Thromboembolism." *British Journal Haematology* 128 (4): 513–9.

Veterans Health Administration. 2010. "VHA Directive 2010-020. 'Anticoagulation Therapy Managementt'" [accessed January 9, 2012]. Available at: www.va.gov/vhapublications/viewpublication.asp?pub_ID=2234.

Wallentin, L., S. Yusuf, M. D. Ezekowitz, M. Alings, M. Flather, M. G. Franzosi, P. Pais, A. Dans, J. Eikelboom, J. Oldgren, J. Pogue, P. A. Reilly, S. Yang, and S. J. Connolly. 2010. "Efficacy and Safety of Dabigatran Compared with Warfarin at Different Levels of International Normalised Ratio Control for Stroke Prevention in Atrial Fibrillation: An Analysis of the RE-LY Trial." *Lancet* 376 (9745): 975–83.

van Walraven, C., A. Jennings, N. Oake, D. Fergusson, and A. J. Forster. 2006. "Effect of Study Setting on Anticoagulation Control: A Systematic Review and Metaregression." *Chest* 129 (5): 1155–66.

White, H. D., M. Gruber, J. Feyzi, S. Kaatz, H. F. Tse, S. Husted, and G. W. Albers. 2007. "Comparison of Outcomes among Patients Randomized to Warfarin

Therapy According to Anticoagulant Control: Results from SPORTIF III and V." *Archives of Internal Medicine* 167 (3): 239–45.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Validating the Patient Safety Indicators in the Veterans Health Administration

Do They Accurately Identify True Safety Events?

Amy K. Rosen, PhD,†‡ Kamal M.F. Itani, MD,‡§‖ Marisa Cevasco, MD, MPH,§¶ Haytham M.A. Kaafarani, MD, MPH,#** Amresh Hanchate, PhD,*‡ Marlena Shin, JD, MPH,* Michael Shwartz, PhD,*†† Susan Loveland, MAT,* Qi Chen, MPH,*† and Ann Borzecki, MD, MPH†‡ ‡‡*

Background: The Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicators (PSIs) use administrative data to detect potentially preventable in-hospital adverse events. However, few studies have determined how accurately the PSIs identify true safety events.

Objectives: We examined the criterion validity, specifically the positive predictive value (PPV), of 12 selected PSIs using clinical data abstracted from the Veterans Health Administration (VA) electronic medical record as the gold standard.

Methods: We identified PSI-flagged cases from 28 representative hospitals by applying the AHRQ PSI software (v.3.1a) to VA fiscal year 2003 to 2007 administrative data. Trained nurse-abstractors used standardized abstraction tools to review a random sample of flagged medical records (112 records per PSI) for the presence of true adverse events. Interrater reliability was assessed. We evaluated PPVs and associated 95% confidence intervals of each PSI and examined false positive (FP) cases to determine why they were incorrectly flagged and gain insight into how each PSI might be improved.

Results: PPVs ranged from 28% (95% CI, 15%−43%) for Postoperative Hip Fracture to 87% (95% CI, 79%−92%) for Postoperative Wound Dehiscence. Common reasons for FPs included conditions that were present on admission (POA), coding errors, and lack of coding specificity. PSIs with the lowest PPVs had the highest proportion of FPs owing to POA.

Conclusions: Overall, PPVs were moderate for most of the PSIs. Implementing POA codes and using more specific ICD-9-CM codes would improve their validity. Our results suggest that additional coding improvements are needed before the PSIs evaluated herein are used for hospital reporting or pay for performance.

Key Words: patient safety, ICD-9-CM coding, criterion validation, adverse events, administrative data

(*Med Care* 2012;50: 74–85)

The Institute of Medicine (IOM) report *To Err is Human* [1] increased national concern and focus on patient safety, prompting demand for evidence-based measures that could be used with administrative data to identify patient safety events. The Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicators (PSIs), released in 2003, were developed in response to this demand and represent a significant contribution to the scientific detection of patient safety events. They are specifically designed to screen for potentially preventable adverse events occurring in the acute inpatient setting. Because they use administrative discharge data (known for its coding variability and inconsistency with respect to diagnoses and procedures), they were viewed as "indicators," rather than as definitive measures, intended for use in quality improvement initiatives, case-finding activities, and monitoring trends.[2]

Although the PSIs underwent a rigorous development process, in which face, content, and construct validity were assessed,[2] and several studies have subsequently evaluated predictive validity,[3–8] concerns about how well the indicators identify true events and accurately reflect hospital performance have heightened with increasing use in public reporting. The National Quality Forum (NQF) recently endorsed 10 PSIs as consensus standards,[9] and the Centers for Medicare & Medicaid Services (CMS) will soon be publicly reporting 6 individual PSIs and a PSI Composite measure on their Hospital Compare website and will be tracking these through their annual payment program.[10,11]

From the *Center for Organization, Leadership and Management Research (COLMR), VA Boston Healthcare System; †Department of Health Policy and Management, Boston University School of Public Health; ‡Boston University School of Medicine; §Department of Surgery, VA Boston Healthcare System; ‖Harvard Medical School; ¶Department of Surgery, Brigham and Women's Hospital; #Department of Surgery, Tufts Medical Center; **Tufts University School of Medicine; ††Department of Operations and Technology Management, Boston University School of Management, Boston, MA; and ‡‡Center for Health Quality, Outcomes and Economic Research, Bedford VAMC, Bedford, MA.

Supported by VA grant HSR&D SDR-07-002.

The authors declare no conflict of interest.

Reprints: Amy K. Rosen, PhD, COLMR, VA Boston Healthcare System, 150 S. Huntington Avenue (152 M), Boston, MA 02130. E-mail: akrosen@bu.edu or amy.rosen2@va.

ISSN: 0025-7079/12/5001-0074

Given these concerns, AHRQ recently conducted the PSI Validation Pilot Project to assess the criterion validity, defined as the degree to which a measure generates data that agree with data from a "gold standard" approach to measuring the same characteristic, of 5 PSIs in a volunteer sample of nonfederal hospitals.[12–16] Validation of additional PSIs, in collaboration with AHRQ, was undertaken by UHC, an alliance of 113 academic medical centers and their affiliated hospitals. Positive predictive value (PPV), the proportion of flagged cases confirmed by chart review to have the PSI event, ranged from 32% for decubitus ulcer to 91% for accidental puncture or laceration.[12–17] These studies highlighted some of the limitations of conducting chart review (eg, unavailability of specific data elements; time intensiveness; poor documentation); they led to recommendations for coding changes to enhance the specificity and PPV of certain indicators[18] and they provided guidelines for use of individual PSIs in quality improvement, public reporting, or pay-for-performance initiatives.

Despite these pilot study results, knowledge about the criterion validity of the PSIs is still relatively limited. Understanding whether AHRQ's results are generalizable to other healthcare settings is particularly critical, as healthcare systems move increasingly toward public reporting of measures. Therefore, we undertook a study to assess the criterion validity of the PSIs in the Veterans Health Administration (VA). The VA is an ideal setting in which to do this. It has a comprehensive, clinically integrated electronic medical record (EMR), which enables centralized access to detailed clinical data and provides a reliable and efficient method for data collection. EMR data can be easily linked with VA discharge data, which are known to have a high level of data element completion and accuracy.[19,20] Further, the VA is strongly committed to making performance and quality data publicly available.[21,22] VA PSI rates will be reported on both the VA and CMS Hospital Compare websites in the near future (personal communication, Dr. Marta Render, 4/11). Finally, although the VA is dedicated to "accountability" through its transparency initiatives, it lacks financial incentives for "upcoding" of diagnoses, and has no penalties or rewards tied to PSI events.

The specific goals of this study were to: (1) assess the criterion validity, specifically the PPV of the PSIs, to determine whether PSI-flagged cases represented true adverse events compared with the gold standard of medical record abstracted data[23]; (2) determine reasons why flagged PSI cases did not represent true events; and (3) recommend modifications to the PSI algorithms to improve PPVs.

METHODS

Study Design and Data

This was a retrospective cross-sectional study using VA administrative and EMR data from VA fiscal year 2003 through 2007 (October 1, 2002 -September 30, 2007). We used hospital discharge information, including demographics, ICD-9-CM coded diagnoses and procedures, and discharge status from the VA's National Patient Care Database Patient Treatment File.[24] On the basis of prior work, we eliminated nonacute care as the PSIs were designed to screen for events in the acute-care setting.[25,26] We also used a previously developed algorithm to distinguish elective from nonelective (ie, urgent/emergent) admissions as VA data lack an admission type and this is required for the denominator of 3 surgical PSIs.[25] VistaWeb was used to access EMR data from all VA facilities.[27]

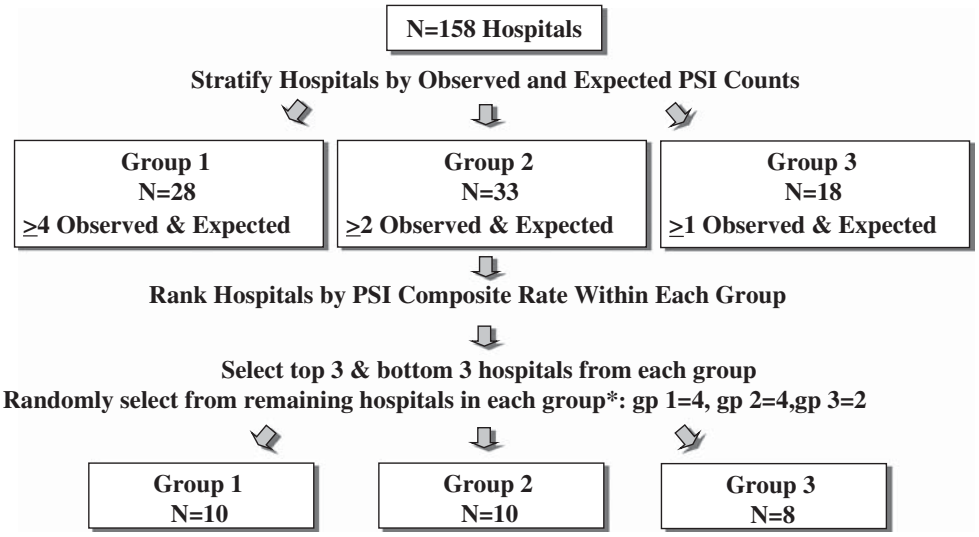
PSI Definitions

The PSIs are constructed as rates or proportions based primarily on relevant ICD-9-CM diagnosis and procedure codes. Each PSI is defined with both a numerator (complication of interest) and denominator (population at risk). We selected 12 of the 20 hospital-level PSIs for study based on their relevance to the VA population, observed VA rates, and their potential preventability. These included decubitus ulcer ("Ulcer"), foreign body left in during procedure ("Foreign Body"), iatrogenic pneumothorax ("Pneumothorax"), central venous catheter-related blood stream infections ("Infections"), postoperative hip fracture ("Fracture"), postoperative hemorrhage or hematoma ("Hemorrhage/Hematoma"), postoperative physiologic and metabolic derangement ("Derangement"), postoperative respiratory failure ("Respiratory Failure"), postoperative pulmonary embolism/deep vein thrombosis ("PE/DVT"), postoperative sepsis ("Sepsis"), postoperative wound dehiscence ("Dehiscence"), and accidental puncture or laceration ("Puncture/Laceration"). (See Appendix for complete definitions.) These PSIs have all been endorsed by the NQF as individual measures.[9] Pneumothorax, Dehiscence, Puncture/Laceration, PE/DVT, and Respiratory Failure will be posted by CMS as individual PSI measures, along with the PSI Composite measure, a combined measure of these 11 PSIs, with the exclusion of Foreign Body [28,29]

We excluded the 4 obstetric PSIs, 2 PSIs with extremely low frequencies (Complications of Anesthesia and Transfusion Reaction), plus 2 PSIs based on mortality (Death in Low Mortality DRGs and Failure to Rescue) because they measure how well hospitals treat complications rather than how well they prevent complications.[30]

Hospital Sampling

Criteria for hospital selection included: (1) obtaining a manageable number of hospitals for chart review; (2) selecting hospitals that represented a spectrum of PSI rates; and (3) ensuring that selected hospitals had an adequate number of PSI events available for chart abstraction. We first applied the AHRQ PSI software (v3.1a) to VA inpatient data to obtain hospital-level PSI counts and PSI composite scores; these were generated by the software.[28,29] Second, we grouped the 158 acute-care VA hospitals into 3 tiers based on their observed and expected PSI counts (Fig. 1). A facility's expected number of PSI events was calculated as the national VA PSI rate multiplied by the PSI denominator of that specific facility. The first tier included hospitals with ≥ 4 observed and expected safety-related events for each PSI, whereas the second and third tiers had ≥ 2 and ≥ 1



Note that Foreign Body and Postoperative Hip Fracture are excluded from the PSIs selected to identify the sample of hospitals. Observed counts are the number of flagged cases per PSI. Expected counts are the number of expected cases per PSI. A facility's expected number of PSI events was calculated as the national VA PSI rate multiplied by the PSI denominator of that specific facility. Our final hospital sample included 28 hospitals (10 in Group 1, 10 in Group 2, and 8 in Group 3). From each of the 28 hospitals, we selected 4 cases per PSI, for a total of 112 cases per PSI.

FIGURE 1. Hospital sampling strategy.

observed and expected events, respectively. Hospitals with < 1 observed or expected safety-related event were excluded, yielding a sample of 79 hospitals. Third, within each tier, we ranked hospitals by the PSI composite measure and included the top 3 and bottom 3 hospitals. Fourth, we randomly selected from the remaining hospitals within each tier to obtain a sample of 28 hospitals. Finally, to assure balanced geographical representation, 3 hospitals were replaced by the next hospital in rank.[31]

Case Identification

Four flagged medical records per PSI were randomly selected from each of the 28 hospitals. On the basis of previously reported PPV estimates, 112 cases per PSI, as available, were selected to ensure reasonably narrow PPV confidence intervals (CI range, 10%−20%).[32]

Medical Record Abstraction

Two trained nurse abstractors reviewed EMRs using standardized data abstraction instruments and guidelines adapted from AHRQ-developed tools[33] or developed de novo if there were no existing tools. Medical records were reviewed for: (1) the occurrence of a safety-related event; (2) demographics, comorbidities, and risk factors; (3) clinical circumstances surrounding the safety-related event; and (4) patient outcomes.

To ensure consistency of abstracted information, we examined interrater reliability (IRR). Two nurses reviewed approximately 10% of the medical records of each PSI. IRR was measured as the percentage of agreement across all questions on each abstraction instrument. Identical records were abstracted in groups of 5 until ≥ 90% agreement was obtained; thereafter, they abstracted different records. Study physicians reviewed questions on which nurses disagreed, with resulting instrument revisions and/or guideline clarifications as appropriate, and also reviewed cases for clarification as needed throughout the abstraction process. Additional IRR assessment was performed on 5 charts toward the end of the abstraction process to check for abstractor reliability drift; subsequent IRR testing revealed ≥ 90% agreement for all PSIs.

Analyses

To determine whether our sample's PSI rates were representative of overall VA hospital PSI rates, we calculated PSI observed and risk-adjusted rates using the PSI software. Next, we categorized flagged cases as either true positives (TPs) or false positives (FPs) based on the application of AHRQ's definition of each PSI and chart-abstracted information. FPs were cases flagged by the PSI software that failed to meet the clinical intent of the indicator. To assess criterion validity, we calculated the PPV of each PSI and associated 95% CIs. PPV was calculated by dividing TPs by the number of flagged cases. We also examined FPs in detail to determine why they were flagged and gain insight into how the PSI might be improved. This included classifying reasons for FP occurrence, such as whether FPs resulted from inappropriate coding or coding limitations. All statistical analyses were performed using SAS version 9.1 (SAS Institute Inc., Cary, NC).

TABLE 1. Rates of Selected Patient Safety Indicators in Our Sample of Hospitals and in All VA Hospitals

PSI No.	Name	Observed Rates in 28 Selected Hospitals	Observed Rates in All VA Hospitals	Risk-adjusted Rates in 28 Selected Hospitals	Risk-adjusted Rates in All VA Hospitals
3	Decubitus Ulcer	15.86	16.41	17.46	17.9
5	Foreign Body Left in During Procedure	0.14	0.12	—+	—+
6	Iatrogenic Pneumothorax	0.68	0.64	0.94	0.90
7	Central Venous Catheter-related Bloodstream Infections	2.21	1.98	2.06	1.89
8	Postoperative Hip Fracture	0.40	0.39	0.61	0.59
9	Postoperative Hemorrhage or Hematoma	3.77	3.89	2.99	3.03
10	Postoperative Physiologic and Metabolic Derangements	2.14	2.15	0.62	0.59
11	Postoperative Respiratory Failure	13.55	14.17	11.06	11.19
12	Postoperative Pulmonary Embolism or Deep Vein Thrombosis	11.43	11.86	10.12	10.39
13	Postoperative Sepsis	6.50	6.67	8.03	7.92
14	Postoperative Wound Dehiscence	6.43	6.18	4.16	3.94
15	Accidental Puncture or Laceration	3.98	3.08	6.61	5.60

Rates are reported per 1000 discharges at risk using PSI software version 3.1a (October 1, 2002-September 30, 2007). Risk adjustment was performed using the AHRQ risk-adjustment software which includes 27 comorbidities, age, sex, and modified DRGs. + PSI 5 is not risk adjusted per PSI software.

RESULTS

Observed PSI rates ranged from 0.14 (Foreign Body) to 15.86 (Ulcer) per 1000 eligible discharges in our 28-hospital sample (Table 1). Ulcer and Respiratory Failure were the most commonly flagged PSIs (15.86 and 13.55, respectively per 1000 discharges). Risk adjustment had a noticeable impact on some of the indicators. Risk-adjusted rates were lower than observed rates for 6 of the PSIs, whereas some risk-adjusted rates were slightly higher than observed, particularly for Ulcer and Sepsis. Given the comparability of observed and risk-adjusted rates between the 2 groups, our hospital sample appeared to be representative of all VA hospitals with respect to PSI rates.

Table 2 displays the PPVs and estimated 95% CIs for selected PSIs and the percentage of cases that were present on admission (POA). PPVs varied considerably, ranging from a low of 28% (95 CI, 15%−43%) for Fracture to a high of 87% (95 CI, 79%−92%) for Dehiscence. Although there were some relatively low PPVs, particularly for Fracture and Ulcer, most PPVs were relatively moderate, ranging from 43% (95% CI, 34−53%) for PE/DVT to 75% (95% CI, 66%−83%) for Hemorrhage/Hematoma. Notably, the PPVs of 2 indicators, Dehiscence and Puncture/Laceration, were relatively high (87% and 85%), respectively, demonstrating good predictive ability.

Many PSIs occurred before the index hospitalization. Overall, 17% of flagged cases were POA, with a range from 0% (Respiratory Failure and Dehiscence) to 59% (Ulcer). POA was also a major reason for flagged cases being deemed as FPs (42% of all FPs were due to POA). For example, among FP cases of Ulcer, 83% were POA; other PSIs with relatively high POA percentages among their FPs included: Fracture (73%), Foreign Body (56%), and Derangement (47%). Because the PSIs with the lowest PPVs also had the highest proportion of FPs owing to POA, we explored the impact on PPV by eliminating POA cases. As expected, the PPVs for these PSIs increased substantially compared to those of other PSIs.

Of the 1266 abstracted cases, 41% were FPs. We categorized reasons for the occurrence of FPs among the 12 PSIs (Table 3). In addition to POA, miscoding and lack of coding specificity were important reasons for mis-flagging cases, accounting for 28% and 16% of FPs, respectively. Individual manuscripts describing the clinical findings of 10 of these PSIs are currently in press.[31,34–40]

We use 1 PSI, PE/DVT, to illustrate the reasons for FPs. Miscoded cases represent those that resulted in a flagged PSI owing to being assigned an incorrect diagnosis or procedure code based on the documentation available in the chart. Among PE/DVT's FPs, there were 24 miscoded cases. These included cases of arterial (not venous) thrombosis, superficial (not deep) vein thrombosis, and cases where a postoperative PE/DVT workup was negative or PE/DVT was a "rule-out" diagnosis. Lack of coding specificity refers to flagged cases that were coded correctly; however, the codes were not specific enough to distinguish between a true PSI event and a nonevent. For example, 14 cases of PE/DVT occurred after admission but before the index procedure. Although these cases were appropriately coded as PE/DVTs, the codes were not specific to "postoperative" events.[31] Two other cases in this category were coded correctly as "thrombophlebitis" events, but because some of the ICD-9-CM codes included in the PSI algorithm do not distinguish superficial phlebitis (eg, inflammation of the superficial veins) from true deep vein thromboses (which is the clinical intent of this PSI), these 2 cases were flagged as PE/DVT events.[31]

There were other less frequent reasons for the occurrence of FPs, some of which were more important with respect to specific PSIs. Three PSIs (Respiratory Failure, Derangement, and Sepsis) required a code for admission type to identify cases that were more likely to be

TABLE 2. Positive Predictive Values and Percentage of Cases That Were Present on Admission Among All Flagged Cases and False Positives by Selected Patient Safety Indicators in the VA

PSI No.	PSI Name	Sample (n)	PPV (%) (95% CI)	POA (%) in All Flagged Cases	POA (%) in False Positives	Sample w/o POA (n)	PPV w/o POA (%) (95% CI)
3	Decubitus Ulcer	112	30 (22–40)	59	83	46	74 (59–86)
5	Foreign Body Left in During Procedure	93	46 (36–55)	30	56	65	66 (53–77)
6	Iatrogenic Pneumothorax	112	73 (64–81)	8	33	103	80 (71–87)
7	Central Venous Catheter-related Bloodstream Infections	112	38 (29–47)	19	30	91	46 (36–57)
8	Postoperative Hip Fracture	46	28 (15–43)	52	73	21	62 (38–82)
9	Postoperative Hemorrhage or Hematoma	112	75 (66–83)	8	32	103	82 (73–89)
10	Postoperative Physiologic and Metabolic Derangements	119	63 (54–72)	18	47	98	77 (67–85)
11	Postoperative Respiratory Failure	112	67 (57–76)	0	0	112	67 (57–76)
12	Postoperative Pulmonary Embolism or Deep Vein Thrombosis	112	43 (34–53)	14	25	96	50 (40–60)
13	Postoperative Sepsis	112	53 (42–64)	14	30	96	61 (51–71)
14	Postoperative Wound Dehiscence	112	87 (79–92)	0	0	112	87 (79–92)
15	Accidental Puncture or Laceration	112	85 (77–91)	5	35	106	90 (82–95)

The number displayed is the number of cases in our sample that were flagged for each of these PSIs. For Postoperative Physiologic and Metabolic Derangements, we flagged 119 cases rather than 112 to ensure that we had an adequate number of diabetes patients who developed abnormalities of glucose control to review in addition to those discharges who developed acute kidney injury requiring dialysis. We were not able to flag 112 cases for Foreign Body and Postoperative Hip Fracture because they were relatively rare events. Parentheses contain 95% confidence intervals (CIs). Positive predictive value (PPV) represents the proportion of true positive cases divided by the number of flagged cases. POA indicates present on admission.

preventable. Misidentification of nonelective admissions as elective accounted for a high percentage of the FP cases for these PSIs, particularly Respiratory Failure and Derangement, (62% and 45% of FPs, respectively).[36,38] A remote or previous history of the PSI before admission was another reason for FPs, particularly for PE/DVT (16% of FPs). Also important for determining if a case was an FP or not was whether the PSI event was related to surgery. Similar to PE/DVT, both Fracture and Sepsis had PSI events that occurred during admission but before surgery (15% and 13%, respectively), and were therefore not related to the index surgery/procedure.

DISCUSSION

This study explored the criterion validity of selected PSIs in the VA. Specifically, we characterized one aspect of criterion validity, PPV, for 12 PSIs that were designed to detect complications that are potentially preventable through good care. Given increasing national focus on improving safety, accurate identification of safety events is critical for guiding improvement efforts. Our overall findings suggest that the original intent of the PSIs, as screening tools to detect potential safety events, may still be the most appropriate use of the PSIs, given the coding inaccuracies and limitations we encountered.

We had several specific study findings. First, consistent with the literature,[26,41–44] we found that PSI rates were generally low; 2 PSIs were relatively rare (Foreign Body and Fracture), with rates less than 1.0 per 1000 discharges. Despite the overall low incidence of individual PSI rates, some PSIs, such as Ulcer (even with POA codes) and Respiratory Failure, had fairly substantial rates, suggesting clinical areas for targeting and opportunities for hospital improvements. As in earlier VA studies, risk adjustment had varying effects on individual PSIs[26,43] in some cases, risk-adjusted rates increased by up to 10%, whereas in others, risk-adjusted rates decreased by up to 35%.

Second, despite evidence on the accuracy and completeness of VA data,[19,20] all of the PSIs misidentified true events to some extent, with considerable PPV variation across PSIs. This variation was likely due to coding inaccuracies or limitations (eg, differences in hospital coding practices, lack of POA codes, lack of precise or meaningful codes, poor documentation). We generally found moderate PPVs for most of the PSIs, ranging from 43% to 75%. Dehiscence had the highest PPV (87%), possibly for 2 reasons: first, wound dehiscence is more likely to occur within the immediate postoperative period compared with other PSIs; and second, in contrast to other PSIs that had high percentages of cases that were POA but were missing POA codes, Dehiscence had no FPs that were due to POA cases.[40]

Our PPV results were generally comparable to those of other non-VA and VA studies[3,12–16,45–47]; for example, for

TABLE 3. Reasons for False Positives by Selected Patient Safety Indicators in the VA

PSI No.	PSI Name	No. FPs	Reasons for Exclusions, n (%)						
			Miscoding (ie, no PSI Event)	Lack of Coding Specificity	POA	Nonelective Admission	History of Event	Occurrence During Admission But Before Surgery	Other
3	Decubitus Ulcer	78	8 (10)		66 (85)		2 (4)		2 (3)
5	Foreign Body Left in During Procedure	50	20 (40)		28 (56)		—		2 (4)
6	Iatrogenic Pneumothorax	30	20 (67)		9 (30)		1 (3)		—
7	Central Venous Catheter-related Bloodstream Infections	70	21 (30)	28 (40)	21 (30)		—		—
8	Postoperative Hip Fracture	33	3 (9)	5 (15)	25 (76)		—	5 (15)	—
9	Postoperative Hemorrhage or Hematoma	28	4 (14)	11 (39)	9 (32)		—	1 (4)	3 (11)
10	Postoperative Physiologic and Metabolic Derangements	44	4 (9)	1 (2)	21 (48)	20 (45)	—	—	3 (7)
11	Postoperative Respiratory Failure	37	7 (19)	4 (11)	—	23 (62)	—	4 (11)	6 (16)
12	Postoperative Pulmonary Embolism or Deep Vein Thrombosis	64	22 (34)	16 (25)	16 (25)		10 (16)	14 (22)	—
13	Postoperative Sepsis	53	12 (23)	13 (25)	16 (30)	12 (23)	—	7 (13)	—
14	Postoperative Wound Dehiscence	15	12 (80)	3 (20)	—		—	—	—
15	Accidental Puncture or Laceration	17	11 (65)	—	6 (35)		—	—	—
Total		519	144	81	217	55	13	31	22

FPs indicates false positive cases; POA, present on admission. The categories of miscoding and nonelective admission are mutually exclusive; other categories may overlap so that the numbers in the column do not necessarily match up with the number of FPs for each PSI; for example, most of the overlap is between cases that occurred during the admission but before surgery and those included under lack of coding specificity. "Other" category includes events that were intentional (eg, a patient with a single kidney underwent a necessary nephrectomy resulting in a need for dialysis and was flagged as PMD), documentation errors (eg, a patient who had diabetic ketoacidosis incorrectly noted as a complication in a discharge summary by a physician, flagged as PMD), postoperative complications that were incidental to surgery (eg, using the PMD example, a patient developed kidney failure from underlying liver problems but it occurred after having a minor surgical procedure, a mouth biopsy under local anesthetic.) These may also overlap with other categories.

PE/DVT, the AHRQ and UHC studies reported PPVs of 47% and 44%, respectively,[15] whereas Henderson et al[45] reported a PPV of 55% compared with our PPV of 43%. A previous VA study compared 5 surgical PSIs with chart-abstracted data obtained from the VA National Surgical Quality Improvement Program (NSQIP), a program designed to promote quality improvement through dissemination of risk-adjusted surgical outcomes to VA facilities.[46] Notably, our PPVs were higher for Sepsis (53% vs. 45%), Dehiscence (87% vs. 72%), and PE/DVT (43% vs. 22%) than those found in the earlier VA study, possibly owing to differences in methodologies and abstraction tools used in the 2 studies, plus definitions of adverse events. Further, that study matched only about 50% of PSI-flagged hospitalizations with NSQIP surgical cases because of NSQIP's exclusion criteria. Nonetheless, there are enough similarities between the various PSI validation studies to suggest that the shortcomings of the PSIs are more inherent to coding-related issues and practices, rather than specific to any 1 study.

Third, there were several reasons for the occurrence of FPs; the 3 most common reasons included diagnoses that were POA, miscoding, and lack of coding specificity. The impact of POA codes on the PSIs is not unique to the VA. A recent AHRQ study found that including POA information reduced the overall number of identified PSI cases by approximately 50%.[48] The effect was especially striking for Ulcer, where as few as 11% of originally flagged cases of Ulcer remained after inclusion of POA codes, followed by Fracture and PE/DVT, where as few as 21% and 43%, respectively, of originally flagged cases remained.[48] The recent CMS mandate to collect POA codes in all Inpatient Prospective Payment System hospitals should help mitigate this issue.[49] Although the VA lags behind the private sector in implementing this policy, it is currently collecting POA codes at the hospital level. Once these codes are incorporated into administrative databases, they will help improve the PPVs of many of the PSIs, particularly those with the highest proportions of FPs owing to POA. Even for PE/DVT, cases that are POA or have a history of PE/DVT would be dropped, potentially increasing the PPV by as much as 13%, albeit decreasing the number of cases flagged.

Another large group of FPs was miscoded cases. As illustrated by PE/DVT, the inability to distinguish the type and nature of the thrombosis was a common reason for miscoding, both in our study and in others.[15,31,45] Lack of coding specificity was also an important reason for FPs among selected PSIs. For the postoperative PSIs, current ICD-9-CM codes do not adequately distinguish the timing of the PSI event (ie, before, during, or after surgery). In addition, the codes may not match up with clinical criteria or nomenclature applied by physicians when documenting patients' diagnoses.[15] For example, at the time of our study, there were no existing codes

that would allow one to distinguish upper extremity from lower extremity DVTs, a distinction that is meaningful for prevention. To remedy some of these specificity issues, AHRQ proposed several coding structure revisions to the ICD-9-CM Coordination and Maintenance Committee, one of which, a new code specifically for upper extremity DVTs, has been approved.[18] Although miscellaneous coding errors and poor physician documentation may still remain as issues, these changes should improve both the codes and coding guidelines for PE/DVT.[14]

General coding changes to improve the inaccuracies and limitations in the ICD-9-CM coding structure, such as implementation of a "timing" code for diagnoses, would help to improve the PPVs of all selected PSIs. Improving the validity of all the PSIs could also be facilitated by standardizing coding practices, such as relying more on evidence-based clinical criteria for diagnostic coding, as well as educating and training medical coders to improve both general and specific coding practices related to adverse event reporting. Providers could assist in this latter process by ensuring that relevant complications are well documented in the discharge summary and that "rule-out" diagnoses are replaced by more precise diagnoses once appropriate tests have been performed. Incorporation of an admission type code denoting an emergent versus elective admission into VA administrative databases would help to improve the PPVs of the 3 PSIs requiring this code in the VA. However, even in UHC member hospitals that code data with respect to admission status, inaccurate admission status coding was still the most frequent reason for FPs,[13] occurring in 5% of Respiratory Failure cases. Nonetheless, these recommendations should be useful in leading to real improvements in quality rather than simply to changes in documentation/coding practices.[13]

Our study has several strengths worth highlighting. Our 28 hospitals were drawn from a nationally representative sample of VA acute-care hospitals. Unlike most other validation studies that relied on paper charts,[12–16,45] we had access to a centralized EMR. This was useful in streamlining the abstraction process and facilitating accurate information retrieval through automated text searching for adverse events and preventing misinterpretation of data owing to poor text legibility. Study clinicians were available to clarify clinical questions throughout the abstraction process, which also helped to maximize the accuracy and reliability of our findings. Finally, we examined agreement between the 2 nurses at different time points, achieving a high level of abstractor agreement at each review.

There were also some limitations. We were not able to report other aspects of criterion validity, namely specificity, sensitivity, or negative predictive value of the PSIs; this was beyond the study's scope and resources. In addition, our sample sizes were small, preventing us from examining whether there was variation in coding accuracy across hospitals.

Although not a specific study goal, we attempted to determine the preventability of the TP cases in our clinically detailed manuscripts.[31,36–40] However, we were unable to determine the actual number of cases associated with quality of care problems given the retrospective nature of the study, lack of evidence-based processes of care for some of the PSIs, information missing from the EMR (such as anesthesia reports), and the lack of control groups. As such, we are limited in concluding whether or not the PSIs, which should in theory be preventable if appropriate perioperative and in-hospital care are provided, do in fact, identify cases that represent true deficiencies in the quality of care and that can lead to or inform meaningful quality improvements.

In conclusion, the PSIs appear to be useful tools for screening, case-finding, and quality improvement. Given the relatively moderate PPVs found, and the need for information on potential preventability and how often true events are missed, we believe it is premature to use the PSIs for public reporting or pay-for-performance. Further, without standardized surveillance systems in place to identify and track safety events nationally, higher PSI rates may be more a marker of improved screening practices or coding differences rather than quality of care.[50] Nonetheless, the PSIs are a step in the right direction, and it is still important to understand what the indicators detect and whether care could be improved.

ACKNOWLEDGMENTS

We are indebted to the following individuals and groups for their assistance in writing the original PSI proposal, participation in the project, guidance and support on all clinical and methodological aspects of the PSIs, and interest in improving the PSIs to make them well-validated and acceptable performance measures: Hillary Mull, MPP, VA Boston Healthcare System; David McClusky, MD, Atlanta VA Medical Center; Andrew Cowan, MD, VA Boston Healthcare System; Kathleen Hickson, RN, MA, VA Boston Healthcare System; Sally MacDonald, RN, Bedford VA Medical Center; Dijon Fasoli, PhD, RN, Bedford VA Medical Center; Julie Lynch, RN, MBA, Bedford VA Medical Center; Kristine Ruggiero, RN, Bedford VA Medical Center; Edward Dunn, MD, ScD, Lexington VA Medical Center; Shibei Zhao, MPH, Bedford VA Medical Center; Arlene Ash, PhD, University of Massachusetts Medical School; Patrick Romano, MD, MPH; UC Davis School of Medicine; Garth Utter, MD, MSc, UC Davis School of Medicine; Patricia Zrelak, PhD, CNRN, CNAA-BC, UC Davis School of Medicine; Jeffrey Geppert, JD, Battelle; Mamatha Pancholi, AHRQ; PSI Validation Study Steering Committee (Leonard Pogach, MD; Anna Alt-White, PhD, RN; William Gunnar, MD; Vickie Montague, RN; Carol Bills, RN, Peter Almenoff, MD; Brian Carney, MD; Allen Gifford, MD; Mary Roseborough, RN); and Joanne Cuny, RN, MBA, AMA.

APPENDIX

Accepted Hospital-Level Indicator Definitions

AHRQ PSI Software Version 3.1 (March, 2007)

Citation: *Patient Safety Indicators Download.* AHRQ Quality Indicators. March 2007. Agency for Healthcare Research and Quality, Rockville, MD. http://www.qualityindicators.ahrq.gov/psi_download.htm.

Indicator	Definition	Numerator	Denominator
PSI 3. Decubitus Ulcer	Cases of decubitus ulcer per 1000 discharges with a length of stay greater than 4 days	Discharges with ICD-9-CM code of decubitus ulcer in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator	All medical and surgical discharges 18 y and older defined by specific DRGs. Exclude cases: With length of stay of less than 5 days With ICD-9-CM code of decubitus ulcer in the principal diagnosis field or in a secondary diagnosis field if present on admission, if known. MDC 9 (Skin, Subcutaneous Tissue, and Breast) MDC 14 (pregnancy, childbirth, and puerperium) with any diagnosis of hemiplegia, paraplegia, or quadriplegia With an ICD-9-CM diagnosis code of spina bifida or anoxic brain damage With an ICD-9-CM procedure code for debridement or pedicle graft before or on the same day as the major operating room procedure (surgical cases only) Admitted from a long-term care facility (SID Admission Source = 3) Transferred from an acute care facility (SID Admission Source = 2)
PSI 5. Foreign Body Left in During Procedure	Discharges with foreign body accidentally left in during procedure per 1000 discharges	Discharges with ICD-9-CM codes for foreign body left in during procedure in any secondary diagnosis field per among cases meeting the inclusion and exclusion rules for the denominator	All medical and surgical discharges, 18 y and older or MDC 14 (pregnancy, childbirth, and puerperium), defined by specific DRGs. Exclude patients with ICD-9-CM codes for foreign body left in during procedure in the principal diagnosis field or secondary diagnosis present on admission, if known
PSI 6. Iatrogenic Pneumothorax	Cases of iatrogenic pneumothorax per 1000 discharges	Discharges with ICD-9-CM code of 512.1 in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator	All medical and surgical discharges 18 y and older defined by specific DRGs. Exclude cases: With ICD-9-CM code of 512.1 in the principal diagnosis field or secondary diagnosis present on admission, if known. MDC 14 (pregnancy, childbirth, and puerperium) With an ICD-9-CM diagnosis code of chest trauma or pleural effusion With an ICD-9-CM procedure code of diaphragmatic surgery repair With any code indicating thoracic surgery or lung or pleural biopsy or assigned to cardiac surgery DRGs
PSI 7. Central Venous Catheter-related Bloodstream Infections* *Patient Safety Indicators Technical Specifications Version 4.2–2010 (Formerly named "Selected Infections Due to Medical Care," Version 3.1)	Cases of ICD-9-CM codes 99662 or 9993 or 99931 per 1000 discharges.	Discharges with selected infections defined by specific ICD-9-CM codes any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator. For discharges before October 1, 2007: *ICD-9-CM Hospital-associated Infection diagnosis codes*: 99662-Due to other vascular device, implant and graft 9993-Other infection.For discharges on or after October 1, 2007: *ICD-9-CM Central Line-associated Bloodstream Infection diagnosis codes*: 99931-Infection due to central venous catheter	All surgical and medical discharges, 18 y and older or MDC 14 (pregnancy, childbirth, and puerperium), defined by specific DRGs or MS-DRGs. Exclude cases: With principal diagnosis of selected infections or secondary diagnosis present on admission With length of stay less than 2 days With any diagnosis or procedure code for immunocompromised state With any diagnosis of cancer
PSI 8. Postoperative Hip Fracture	Cases of in-hospital hip fracture per 1000 surgical discharges with an operating room procedure.	Discharges with ICD-9-CM code for hip fracture in any secondary diagnosis of field among cases meeting the inclusion and exclusion rules for the denominator.	All surgical discharges 18 y and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure.

(*continued*)

TABLE. (*continued*)

Indicator	Definition	Numerator	Denominator
			Exclude cases: With ICD-9-CM code for hip fracture in the principal diagnosis field or secondary diagnosis present on admission, if known Where the only operating procedure is hip fracture repair Where a procedure for hip fracture repair occurs before or on the same day as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available* With diseases and disorders of the musculoskeletal system and connective tissue (MDC 8) With principal diagnosis (or secondary diagnosis present on admission, if known) of seizure, syncope, stroke, coma, cardiac arrest, poisoning, trauma, delirium and other psychoses, or anoxic brain injury With any diagnosis of metastatic cancer, lymphoid malignancy or bone malignancy, or self-inflicted injury MDC 14 (pregnancy, childbirth, and puerperium)
PSI 9. Postoperative Hemorrhage or Hematoma	Cases of hematoma or hemorrhage requiring a procedure per 1000 surgical discharges with an operating room procedure.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with either of the following: ICD-9-CM codes for postoperative hemorrhage in any secondary diagnosis field and a code for drainage of hematoma in any procedure code field ICD-9-CM codes for postoperative hematoma in any secondary diagnosis field and a code for postoperative control of hemorrhage in any procedure code field.	All surgical discharges 18 y and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: With preexisting condition (principal diagnosis or secondary diagnosis present on admission, if known) of postoperative hemorrhage or postoperative hematoma Where the only operating room procedure is postoperative control of hemorrhage or drainage of hematoma Where a procedure for postoperative control of hemorrhage or drainage of hematoma occurs before the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* MDC 14 (pregnancy, childbirth and the puerperium)
PSI 10. Postoperative Physiologic and Metabolic Derangements	Cases of specified physiological or metabolic derangement per 1000 elective surgical discharges with an operating room procedure.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for physiologic and metabolic derangements in any secondary diagnosis field. Discharges with acute renal failure (subgroup of physiologic and metabolic derangements) must be accompanied by a procedure code for dialysis (3995, 5498).	All elective* surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. *Defined by admit type. Exclude cases: With preexisting condition (principal diagnosis or secondary diagnosis present on admission, if known) of physiologic and metabolic derangements or chronic renal failure With acute renal failure where a procedure for dialysis occurs before or on the same day as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be*

(*continued*)

TABLE. (*continued*)

Indicator	Definition	Numerator	Denominator
			slightly lower than if the information was available With both a diagnosis code of ketoacidosis, hyperosmolarity, or other coma (subgroups of physiologic and metabolic derangements coding) and a principal diagnosis of diabetes With both a secondary diagnosis code for acute renal failure (subgroup of physiologic and metabolic derangements coding) and a principal diagnosis of acute myocardial infarction, cardiac arrhythmia, cardiac arrest, shock, hemorrhage, or gastrointestinal hemorrhage MDC 14 (pregnancy, childbirth and puerperium)
PSI 11. Postoperative Respiratory Failure	Cases of acute respiratory failure per 1000 elective surgical discharges with an operating room procedure.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for acute respiratory failure (518.81) in any secondary diagnosis field (After 1999, include 518.84). OR Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for reintubation procedure as follows: (96.04) One or more days after the major operating room procedure code (96.70 or 97.71) Two or more days after the major operating room procedure code (96.72) zero or more days after the major operating room procedure code.	All elective* surgical discharges age 18 and over defined by specific DRGs and an ICD-9-CM code for an operating room procedure. *Defined by admit type. Exclude cases: With preexisting (principal diagnosis or secondary diagnosis present on admission, if known) acute respiratory failure With an ICD-9-CM diagnosis code of neuromuscular disorder Where a procedure for tracheostomy is the only operating room procedure or tracheostomy occurs before the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* MDC 14 (pregnancy, childbirth, and puerperium) MDC 4 (diseases/disorders of respiratory system) MDC 5 (diseases/disorders of circulatory system)
PSI 12. Postoperative Pulmonary Embolism or Deep Vein Thrombosis	Cases of deep vein thrombosis (DVT) or pulmonary embolism (PE) per 1000 surgical discharges with an operating room procedure.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for deep vein thrombosis or pulmonary embolism in any secondary diagnosis field.	All surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: With preexisting (principal diagnosis or secondary diagnosis present on admission, if known) deep vein thrombosis or pulmonary embolism where a procedure for interruption of vena cava is the only operating room procedure Where a procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* MDC 14 (pregnancy, childbirth and puerperium)
PSI 13. Postoperative Sepsis	Cases of sepsis per 1000 elective surgery patients with an operating room procedure and a length of stay of 4 days or more.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code for sepsis in any secondary diagnosis.	All elective* surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure.

(*continued*)

TABLE. (*continued*)

Indicator	Definition	Numerator	Denominator
			*Elective-Admission type # is recorded as elective (Admission Type = 3). Exclude cases: With preexisting (principal diagnosis or secondary present on admission, if known) sepsis or infection With any code for immunocompromised state or cancer MDC 14 (pregnancy, childbirth, and puerperium) With a length of stay of less than 4 days
PSI 14. Postoperative Wound Dehiscence	Cases of reclosure of postoperative disruption of abdominal wall per 1000 cases of abdominopelvic surgery.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code for reclosure of postoperative disruption of abdominal wall (54.61) in any procedure field.	All abdominopelvic surgical discharges age 18 and older. Exclude cases: Where a procedure for reclosure of postoperative disruption of abdominal wall occurs before or on the same day as the first abdominopelvic surgery procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available* Where length of stay is less than 2 days With immunocompromised state MDC 14 (pregnancy, childbirth, and puerperium)
PSI 15. Accidental Puncture or Laceration	Cases of technical difficulty (eg, accidental cut or laceration during procedure) per 1000 discharges.	Discharges among cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code denoting technical difficulty (eg, accidental cut, puncture, perforation, or laceration) in any secondary diagnosis field.	All medical and surgical discharges age 18 y and older defined by specific DRGs. Exclude cases: With ICD-9-CM code denoting technical difficulty (eg, accidental cut, puncture, perforation, or laceration) in the principal diagnosis field or secondary diagnosis present on admission, if known MDC 14 (pregnancy, childbirth, and puerperium)

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS, eds. To err is human: building a safer health system. A report of the Committee on Quality of Health Care in America, Institute of Medicine. Washington, DC: National Academy Press; 2000.
2. Agency for Healthcare Research and Quality. AHRQ Quality Indicators: Guide to Patient Safety Indicators. Rockville, MD; March 2003.
3. Gallagher B, Cen L, Hannan E. Validation of AHRQ's patient safety indicator for accidental puncture or laceration. In: Henriksen K, Battles J, Marks E, et al, eds. Advances in Patient Safety: From Research to Implementation. Concepts and Methodology. Vol 2. Rockville, MD: Agency for Healthcare Research and Quality; 2005. AHRQ Publication No. 05-0021-2.
4. Grobman WA, Feinglass J, Murthy S. Are the Agency for Healthcare Research and Quality obstetric trauma indicators valid measures of hospital safety? *Am J Obstet Gynecol*. 2006;195:868–874.
5. Haller G, Myles PS, Wolfe R, et al. Validity of unplanned admission to an intensive care unit as a measure of patient safety in surgical patients. *Anesthesiology*. 2005;103:1121–1129.
6. Rivard P, Luther S, Christiansen C, et al. Using Patient Safety Indicators to estimate the impact of potential adverse events on outcomes. *Med Care Res Rev*. 2008;65:67–87.
7. Shufelt JL, Hannan EL, Gallagher BK. The Postoperative Hemorrhage and Hematoma Patient Safety Indicator and its risk factors. *Am J Med Qual*. 2005;20:210–218.
8. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA*. 2003;290:1868–1874.
9. National Quality Forum. NQF Endorsed Standards. Available at: http://www.qualityforum.org/Measures_List.aspx. Accessed June 1, 2010.
10. Centers for Medicare and Medicaid Services (CMS). CMS Hospital Compare Website. Available at: http://www.cms.gov/HospitalQualityInits/11_HospitalCompare.asp2010. Accessed June 1, 2010.
11. Reporting Hospital Quality Data for Annual Payment Update. Hospital Inpatient Quality Reporting (IQR) Program Overview. Available at: http://www.qualitynet.org/dcs/ContentServer?cid = 1138115987129&pagename = QnetPublic%2FPage%2FQnetTier2&c = Page. Accessed December 10, 2010.
12. Sadeghi B, Baron R, Zrelak P, et al. Cases of iatrogenic pneumothorax can be identified from ICD-9-CM coded data. *Am J Med Qual*. 2010;25:218–224.
13. Utter GH, Cuny J, Sama P, et al. Detection of Postoperative Respiratory Failure: How predictive is the Agency for Healthcare Research and Quality's Patient Safety Indicator? *J Am Coll Surg*. 2010;211:347–354.
14. Utter GH, Zrelak PA, Baron R, et al. Positive predictive value of the AHRQ Accidental Puncture or Laceration Patient Safety Indicator. *Ann Surg*. 2009;250:1041–1045.
15. White R, Sadeghi B, Tancredi D, et al. How valid is the ICD-9-CM based AHRQ Patient Safety Indicator for Postoperative Venous Thromboembolism? *Med Care*. 2009;47:1237–1243.

16. Zrelak P, Sadeghi B, Utter G, et al. Positive predictive value of the Agency for Healthcare Research and Quality Patient Safety Indicator for Central Line-Related Bloodstream Infection ("Selected Infections due to Medical Care"). *J Health Qual*. 2011;1–8.
17. Zaratkiewicz S. Pressure Ulcer 2008 Benchmarking Project Findings and Conclusions. Knowledge Transfer Meeting. UHC 2009.
18. Centers for Disease Control and Prevention. ICD-9-CM Coordination and Maintenance Committee Meeting, Diagnosis Agenda, September 24–25, 2008. Available at: http://www.cdc.gov/nchs/data/icd9/agenda-Sep08.pdf. Accessed December 15, 2010.
19. Singh JA. Accuracy of Veterans Affairs databases for diagnoses of chronic diseases. *Prev Chronic Dis*. 2009;6:A126.
20. Kashner TM. Agreement between administrative files and written medical records: A case of the Department of Veterans Affairs. *Med Care*. 1998;36:1324–1336.
21. Department of Veterans Affairs. Secretary Shinseki Releases Hospital Report Card. VA Office of Public Affairs Media Relations. Available at: http://www1.va.gov/opa/pressrel/pressrelease.cfm?id=1830. Accessed November 15, 2010.
22. Department of Veterans Affairs. Open Government Plan. June 2010. Prepared on behalf of Eric K. Shinseki, Secretary. Accessed January 15, 2011.
23. Murff HJ, Patel VL, Hripcsak G, et al. Detecting adverse events for patient safety research: A review of current methodologies. *J Biomed Inform*. 2003;36:131–143.
24. VA Information Resource Center. VIReC Research User Guide: FY2002 VHA Medical SAS Inpatient Datasets. Available at: http://www.virec.research.va.gov/References/RUG/RUG-Inpatient02.pdf. Accessed December 1, 2009.
25. Rivard P, Elwy R, Loveland S, et al, eds. *Applying Patient Safety Indicators (PSIs) Across Health Care Systems: Achieving Data Comparability Vol 2: 7–25*. Rockland, MD: Agency for Healthcare Research and Quality; 2005.
26. Rosen AK, Rivard P, Zhao S, et al. Evaluating the patient safety indicators: How well do they perform on Veterans Health Administration data? *Med Care*. 2005;43:873–884.
27. Veterans Health Administration. VistA Web Overview. Available at: http://vista.med.va.gov/vistaweb. Accessed January 10, 2010.
28. Agency for Healthcare Research and Quality. Patient Safety Indicators Software SAS Documentation. [Version 3.2]. March 2003.
29. Agency for Healthcare Research and Quality. AHRQ Quality Indicators: Patient Safety Indicator Composite SAS Documentation. [Version 3.1a]. March 2007.
30. Rosen A, Loveland S, Romano P, et al. Effects of resident duty hour reform on surgical and procedural Patient Safety Indicators among hospitalized Veterans Health Administration and Medicare patients. *Med Care*. 2009;47:723–731.
31. Kaafarani HM, Borzecki AM, Itani KM, et al. Validity of Selected Patient Safety Indicators: Opportunities and concerns. *J Am Coll Surg*. 2011;212:924–934.
32. Lawthers AG, McCarthy EP, Davis RB, et al. Identification of in-hospital complications from claims data. *Is it valid? Med Care*. 2000;38:785–795.
33. AHRQ Quality Indicators Validation Pilot. Available at: http://qualityindicators.ahrq.gov/validationpilot.htm. Accessed March 1, 2010.
34. Cevasco M, Borzecki A, O'Brien W, et al. Validity of the AHRQ Patient Safety Indicator #7: Central Venous Catheter Related Bloodstream Infections. *J Am Coll Surg*. 2011;212:984–990.
35. Chen Q, Rosen A, Cevasco M, et al. Detecting Patient Safety Indicators: How valid is Foreign Body Left During Procedure in the Veterans Health Administration? *J Am Coll Surg*. 2011;212:977–983.
36. Borzecki A, Cevasco M, Chen Q, et al. How valid is the AHRQ Patient Safety Indicator Postoperative Physiologic and Metabolic Derangement? *J Am Coll Surg*. 2011;212:968–976.
37. Borzecki A, Kaafarani H, Cevasco M, et al. How valid is the AHRQ Patient Safety Indicator Postoperative Hemorrhage or Hematoma? *J Am Coll Surg*. 2011;212:946–953.
38. Borzecki A, Kaafarani H, Utter G, et al. How valid is the AHRQ Patient Safety Indicator Postoperative Respiratory Failure? The Veterans Health Administration experience. *J Am Coll Surg*. 2011;212:935–945.
39. Cevasco M, Borzecki A, Chen Q, et al. Positive predictive value of the AHRQ Patient Safety Indicator Postoperative Sepsis. *J Am Coll Surg*. 2011;212:954–961.
40. Cevasco M, Borzecki A, McCluskey D, et al. Positive predictive value of the AHRQ Patient Safety Indicator Postoperative Wound Dehiscence. *J Am Coll Surg*. 2011;212:962–967.
41. Rivard P, Elixhauser A, Christiansen C, et al. Testing the association between Patient Safety Indicators and hospital structural characteristics in VA and nonfederal hospitals. *Med Care Res Rev*. 2010;67:321–341.
42. Romano PS, Geppert JJ, Davies S, et al. A national profile of patient safety in U.S. hospitals. *Health Affairs*. 2003;22:154–166.
43. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: Lessons from the Veterans Administration. *Med Care*. 2006;44:850–861.
44. Scanlon MC, Harris JM 2nd, Levy F, et al. Evaluation of the agency for healthcare research and quality pediatric quality indicators. *Pediatrics*. 2008;121:e1723–e1731.
45. Henderson K, Recktenwald A, Reichley R, et al. Clinical validation of the AHRQ Postoperative Venous Thromboembolism Patient Safety Indicator. *Jt Comm J Qual Patient Saf*. 2009;35:370–376.
46. Romano P, Mull H, Rivard P, et al. Validity of selected AHRQ Patient Safety Indicators based on VA National Surgical Quality Improvement Program data. *Health Serv Res*. 2009;44:182–204.
47. Zhan C, Battles J, Chiang Y, et al. The validity of ICD-9-CM codes in identifying Postoperative Deep Vein Thrombosis and Pulmonary Embolism. *Jt Comm J Qual Patient Saf*. 2007;33:326–331.
48. Houchens R, Elixhauser A, Romano P. How often are potential patient safety events present on admission? *Jt Comm J Qual Patient Saf*. 2008;34:154–163.
49. Department of Health and Human Services (DHHS). Hospital-Acquired Conditions (HAC) in Acute Inpatient Prospective Payment System (IPPS) Hospitals. May 2010. Available at: www.cms.gov/hospitalacqcond. Accessed November, 2010.
50. Haut ER, Noll K, Efron DT, et al. Can increased incidence of Deep Vein Thrombosis (DVT) be used as a marker of quality of care in the absence of standardized screening? The potential effect of surveillance bias on reported DVT rates after trauma. *J Trauma*. 2007; 63:1132–1135.

Examining the Impact of the AHRQ Patient Safety Indicators (PSIs) on the Veterans Health Administration

The Case of Readmissions

Amy K. Rosen, PhD,† Susan Loveland, MAT,* Marlena Shin, JD, MPH,* Michael Shwartz, PhD,*‡ Amresh Hanchate, PhD,*† Qi Chen, MPH,*§ Haytham M.A. Kaafarani, MD, MPH,‖¶ and Ann Borzecki, MD, MPH†§#*

Background: By focusing primarily on outcomes in the inpatient setting one may overlook serious adverse events that may occur after discharge (eg, readmissions, mortality) as well as opportunities for improving outpatient care.

Objective: Our overall objective was to examine whether experiencing an Agency for Healthcare Research and Quality Patient Safety Indicator (PSI) event in an index medical or surgical hospitalization increased the likelihood of readmission.

Methods: We applied the Agency for Healthcare Research and Quality PSI software (version 4.1.a) to 2003–2007 Veterans Health Administration inpatient discharge data to generate risk-adjusted PSI rates for 9 individual PSIs and 4 aggregate PSI measures: any PSI event and composite PSIs reflecting "Technical Care," "Continuity of Care," and both surgical and medical care (Mixed). We estimated separate logistic regression models to predict the likelihood of 30-day readmission for individual PSIs, any PSI event, and the 3 composites, adjusting for age, sex, comorbidities, and the occurrence of other PSI(s).

Results: The odds of readmission were 23% higher for index hospitalizations with any PSI event compared with those with no event [confidence interval (CI), 1.19–1.26], and ranged from 22% higher for Iatrogenic Pneumothorax (CI, 1.03–1.45) to 61% higher for Postoperative Wound Dehiscence (CI, 1.27–2.05). For the composites, the odds of readmission ranged from 15% higher for the Technical Care composite (CI, 1.08–1.22) to 37% higher for the Continuity of Care composite (CI, 1.26–1.50).

Conclusions: Our results suggest that interventions that focus on minimizing preventable inpatient safety events as well as improving coordination of care between and across settings may decrease the likelihood of readmission.

Key Words: patient safety, readmissions, adverse events, coordination of care

(*Med Care* 2013;51: 37–44)

Despite improvements in patient safety resulting from the Institute of Medicine report "To Err is Human,"[1] shortcomings in patient safety remain a major concern. Only 36% of safety-related hospital measures (eg, postoperative complications) improved at a rate >5% per year, compared with 84% of non–safety-related measures (eg, measures of effectiveness, timeliness, and patient centeredness).[2] Although tracking safety event trends is complicated, standardized measures such as the Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicators (PSIs) have helped identify hospitalizations with potentially preventable adverse events (AEs), hospital variation in the safety of care,[3–6] and the consequences of AEs [eg, excess mortality, length of stay (LOS), and costs] during hospitalization.[7–9]

Nevertheless, by focusing excessively on outcomes in the inpatient setting one may overlook serious AEs among inpatients that may manifest after discharge,[10,11] important postdischarge safety events,[12,13] and opportunities for improvement in postdischarge patient care.[11,14] Recent studies have expanded on earlier work linking PSI events with in-hospital adverse outcomes,[7–9] attributing postdischarge outcomes, such as readmissions, to PSI events that occurred during the index admission.[10,11] In 1 study of adult surgery patients treated in nonfederal hospitals, the 30-day readmission rate increased by 44% when a PSI event occurred during the index admission.[11] Another study of privately insured adult surgery patients suggested that PSI events in the index admission were responsible for 11% of all deaths and 2% of all readmissions at 90 days.[10]

From the *Center for Leadership, Organization and Management Research (COLMR), VA Boston Healthcare System; †Boston University School of Medicine; ‡Department of Operations and Technology Management, Boston University School of Management; §Department of Health Policy and Management, Boston University School of Public Health; ‖Division of Trauma, Emergency Surgery and Surgical Critical Care, Massachusetts General Hospital; ¶Harvard Medical School, Boston; and #Center for Health Quality, Outcomes and Economic Research, Bedford VAMC, Bedford, MA.

Supported by VA grant HSR&D SDR-07-002.

The authors declare no conflict of interest.

Reprints: Amy K. Rosen, PhD, Center for Leadership, Organization and Management Research, VA Boston Healthcare System, Boston, MA 02130. E-mail: akrosen@bu.edu or amy.rosen2@va.gov.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website, www.lww-medicalcare.com.

ISSN: 0025-7079/13/5101-0037

We build on the sparse literature in this area by examining whether the occurrence of a PSI event in the index hospitalization increases the likelihood of readmission in the Veterans Health Administration (VA). The VA is the nation's largest integrated health care system, currently offering a full continuum of care to approximately 8 million veterans.[15] We also expand the previous focus from surgical cases only to both medical and surgical cases, making our findings particularly timely given the recent posting of VA 30-day readmission rates for pneumonia, heart failure (HF), and acute myocardial infarction on the Centers for Medicare and Medicaid Services Hospital Compare website.[16,17] Readmissions have emerged nationally as an important outcome measure, as they occur frequently (all-cause readmission rates range from 13% to 20%),[14,17] account for substantial costs (estimates as high as $25 billion per year have been reported),[18] and reflect poor quality of care during the index hospitalization, breakdown in continuity of care, poor discharge planning, or the occurrence of an AE in the index hospitalization.[11,14,19]

Although a recent systematic review of the literature on readmission prediction models suggests that patient-level factors, such as sociodemographic and clinical variables, are strongly associated with readmission risk, it also noted that the lack of hospital-level and system-level factors (including the quality of inpatient care) limits the models' ability to predict readmission.[20] By examining whether a PSI event in the index hospitalization increases the likelihood of readmission, we may be able to improve the ability of the models to target those at high risk of readmission and thus facilitate the design of transitional care interventions.

Our specific objectives were to: (1) compare demographic and clinical characteristics of index hospitalizations with and without PSIs; (2) examine 30-day "all-cause" readmission rates for index hospitalizations with and without PSIs; (3) assess whether experiencing a PSI event in the index hospitalization affects the reason for readmission; and (4) examine the likelihood of 30-day readmission on the basis of the type of PSI event that occurred in the index hospitalization.

METHODS

Data Sources

Our primary source of inpatient data was the VA Patient Treatment File, an administrative database that contains information on all veterans discharged from VA acute-care facilities, such as demographics, principal and secondary diagnoses, and surgical procedures.[21] Because the PSIs were designed to screen for patient safety events in acute-care hospitals, as in a previous work, we eliminated the nonacute portion of care, creating a hospital discharge file containing only acute care.[22] We linked the Patient Treatment File data to dates of death from the VA Vital Status File to obtain information on deaths within 30 days after discharge. Study approvals were received from the relevant Institutional Review Boards.

Sample

Our initial sample (n = 2,332,794) comprised all acute-care discharges from VA hospitals from fiscal years 2003 to 2007 (October 1, 2002 through September 30, 2007). Similar to the Centers for Medicare and Medicaid Services' Hospital Compare methodology, we combined hospitalizations in which the subsequent admission was on the same or next day after the prior discharge.[23] There were a total of 60,032 admissions that were combined with a preceding discharge, which reduced the final sample to 2,272,762 discharges. For these aggregated hospitalizations, admission attributes (eg, principal diagnosis) came from the first hospitalization and discharge attributes (eg, discharge location) came from the last hospitalization.

We defined an index hospitalization as the first hospitalization that occurred on or after October 1, 2002 (the first day of the study period) or 31 days or more after the index discharge. Thus, patients could have multiple index hospitalizations (it is noteworthy that 55% had only 1 index hospitalization over the 5 y). Consistent with other studies[11,14,23] we excluded index hospitalizations in which the patient left against medical advice (n = 32,980) or was transferred to a non-VA hospital (n = 14,807). We also excluded index hospitalizations documented as occurring after the patient's date of death (n = 3,899) (ie, data errors) or discharged after August 9, 2007 (n = 60,809), ensuring that any index hospitalization occurring toward the end of our data would have a large enough window after discharge to capture a readmission. This cutoff date, 51 days before the last day of the fiscal year 2007, accounted for the 30-day span from discharge to potential readmission as well as the readmission's LOS (95% of all hospitalizations had a LOS <21 d).

As final steps, we eliminated index hospitalizations that died in-hospital (n = 58,767), and excluded 290 hospitalizations that did not meet eligibility criteria for any of the PSIs used in the study (eg, obstetric discharges). We also retained the readmission information associated with all index hospitalizations to examine reasons for readmission. However, for hospitalizations with multiple readmissions occurring <30 days apart, we kept only the first readmission, excluding 31,696 subsequent readmissions.

This yielded a final sample of 1,807,488 index hospitalizations and 262,026 readmissions for study. We divided the index hospitalizations into 2 groups for analytical purposes: index hospitalizations with a PSI event(s) (n = 27,958) and those without a PSI event(s) (n = 1,779,530).

The PSIs

The PSIs were designed specifically to capture potentially preventable events that represent compromised patient safety in the acute-care setting, such as complications after surgeries, procedures, or medical care. The original 20 hospital-level PSIs resulted from a 4-step process including literature review, evaluation of candidate PSIs by clinical panels, consultation with coding experts, and empirical analyses.[24] The PSI software uses secondary diagnoses, procedures, and other hospital discharge information to flag hospitalizations for potential safety-related events. The PSIs are calculated as rates, defined with a numerator (complication of interest) and denominator (population at risk). Currently there are 17 PSIs, 7 relevant to surgical discharges

only, 8 to both medical and surgical discharges, and 2 to obstetric discharges.[25]

We excluded the following PSIs from all analyses: 2 obstetric PSIs as they are not relevant to the VA population; 2 PSIs related to in-hospital mortality: PSI #2 (Death in Low Mortality DRGs) and PSI #4 (Complications in Surgical Inpatients); and PSI #16 (Transfusion Reaction) which is very rare. Three PSIs, PSI #8 (Postoperative Hip Fracture), PSI #3 (Pressure Ulcer), and PSI #5 (Foreign Body Left in During Procedure), were used in aggregated PSI measures only (described below)—the former 2 because they are present on admission about 70% of the time[26] and therefore are not associated with care in the index hospitalization, and the latter because it is relatively rare.[27] This yielded 9 PSIs for individual analyses (see Table, Supplemental Digital Content 1, http://links.lww.com/MLR/A354 for PSI definitions).

Four aggregated PSI measures were used for analyses. The first measure was designed to capture "any PSI event" (yes/no) and contained the 9 individual PSIs plus PSIs #3, #5, and #8. Three PSI composite measures, developed in previous work to compensate for the relatively low rates of individual PSIs, were also included in analyses.[28] These measures, empirically derived through factor analyses and clinical judgment, grouped individual PSIs into clinically meaningful categories. The "Continuity of Care" composite contained PSIs that reflect continuity of care and medical management of high-risk patients in the perioperative setting: PSI #10 (Postoperative Physiologic and Metabolic Derangements), PSI #11 (Postoperative Respiratory Failure), and PSI #13 (Postoperative Sepsis). The "Technical Care" composite included PSIs reflecting technical skill–based care: PSI #5 (Foreign Body Left in During Procedure), PSI #9 (Postoperative Hemorrhage or Hematoma), PSI #14 (Postoperative Wound Dehiscence), and PSI #15 (Accidental Puncture or Laceration). The "Mixed" composite included both surgical and medical PSIs: PSI #6 (Iatrogenic Pneumothorax) and PSI #7 (Central Venous Catheter-related Bloodstream Infections), PSI #8 (Postoperative Hip Fracture), and PSI #12 (Postoperative Pulmonary Embolism/Deep Vein Thrombosis).

Because the PSIs were developed and tested using hospital discharge abstracts from AHRQ's HCUP State Inpatient Databases, they require UB-92 (1992 Uniform Bill) data elements from hospital claims,[22] some of which are missing from VA databases. Consequently, we used previously developed algorithms to calculate certain variables, such as principal procedure, that were not readily available from the data.[29]

Outcome Measure

The primary outcome measure was a 30-day all-cause readmission, defined as a readmission occurring for any diagnosis within 30 days of an index discharge.

Analyses

Analyses were performed by applying the PSI software (version 4.1a) and the statistical analysis system (SAS, version 9.1.3) to our dataset. We ran descriptive statistics, *t* tests, and χ^2 tests as appropriate, to compare demographic and clinical characteristics, as well as rates of readmission, between index hospitalizations with and without any PSI event. We also explored whether having any PSI event in the index hospitalization (prior PSI event) was related to the reason for readmission, defined by the principal diagnosis or the primary cause of readmission.[21] To provide a more meaningful description of the reason for each readmission, we grouped the principal diagnosis of the readmission associated with each index hospitalization into a discrete category using the clinical classification software (CCS), a tool developed by AHRQ for clustering patient diagnoses and procedures into a manageable number of clinically meaningful categories.[30]

As a final step, we developed separate logistic regression models to predict the likelihood of 30-day readmission for any PSI event, selected PSIs, and each of the 3 PSI composite measures, controlling for age, sex, 27 comorbidities that are included in the AHRQ comorbidity software,[31] and the occurrence of any other PSI(s). To explore the impact of multiple PSIs, we ran a logistic regression model among hospitalizations with at least 1 PSI, and included a variable indicating 2 or more PSIs. Odds ratios and 95% confidence intervals (CIs) were calculated for each of the 13 models. As a sensitivity analysis to account for the potential correlation among repeated hospitalizations for the same patient, we reran the logistic regression models nesting hospitalizations within patients; to account for any changes in readmissions over time, we included dummy variables for "year."

RESULTS

Characteristics of Index Hospitalizations

As shown in Table 1, our sample was predominantly male, with a mean age of 64.8 years (SD = 12.9) and LOS of 6.6 days (SD = 35.4). Compared with index hospitalizations without any PSI event, those with any PSI event were slightly older (67.8 vs. 64.8 y, respectively), had much longer LOS (24.2 vs. 6.3 d), were more likely to be hospitalized for a surgical rather than a medical inpatient stay (62.1% vs. 37.9%), and were more likely to die within 30 days after discharge (9.4% vs. 3.4%). Hospitalizations with any PSI event also had higher rates of HF and renal failure compared with hospitalizations without events (10.4% vs. 6.9% and 8.1% vs. 6.1%, respectively). In addition, readmission rates were higher for index hospitalizations with any PSI event compared with those without an event (18.6% vs. 14.4%, respectively) ($P < 0.0001$ for all listed comparisons).

Rates of readmission were consistently higher for index hospitalizations with selected PSIs and PSI composites compared with hospitalizations without these events (Table 2). Differences in readmission rates were greatest for index hospitalizations with and without Postoperative Physiologic and Metabolic Derangements (23.8% vs. 11.4%, respectively). Among composites, the largest difference in readmission rates was for index hospitalizations with and without the Continuity of Care composite (18.7% vs.

TABLE 1. Selected Demographic and Clinical Characteristics of Index Hospitalizations Discharged Alive With and Without Any PSI Event (Fiscal Years 2003–2007)*

	Total Index Hospitalizations	Index Hospitalizations With Any PSI Event†	Index Hospitalizations Without Any PSI Event
N (%)	1,807,488 (100%)	27,958 (2%)	1,779,530 (98%)
Age, mean (SD)	64.8 (12.9)	67.8 (12.2)	64.8 (12.9)
Male (%)	96.4%	97.2%	96.4%
Length of stay (d), mean (SD)	6.6 (35.4)	24.2 (48.7)	6.3 (35.1)
Death within 30 d after discharge, N (%)	63,452 (3.5%)	2613 (9.4%)	60,839 (3.4%)
No. comorbidities, mean (SD)	1.9 (1.4)	2.0 (1.4)	1.9 (1.4)
Diabetes, N (%)	517,485 (28.6%)	8020 (28.7%)	509,465 (28.6%)
Heart failure, N (%)	125,857 (7.0%)	2901 (10.4%)	122,956 (6.9%)
Renal failure, N (%)	110,273 (6.1%)	2255 (8.1%)	108,018 (6.1%)
Valvular disease, N (%)	37,759 (2.1%)	832 (3.0%)	36,927 (2.1%)
Hypertension, N (%)	952,525 (52.7%)	12,766 (45.7%)	939,759 (52.8%)
Chronic lung diseases, N (%)	331,210 (18.3%)	5602 (20.0%)	325,608 (18.3%)
Obesity, N (%)	75,902 (4.2%)	803 (2.9%)	75,099 (4.2%)
Paralysis, N (%)	46,614 (2.6%)	809 (2.9%)	45,805 (2.6%)
Liver disease, N (%)	75,650 (4.2%)	1064 (3.8%)	74,586 (4.2%)
Alcohol/drug abuse, N (%)	161,347 (8.9%)	1958 (7.0%)	159,389 (9.0%)
30-day readmission,‡ N (%)	262,026 (14.5%)	5193 (18.6%)	256,833 (14.4%)
Medical DRGs, N (%)	1,367,079 (75.6%)	10,599 (37.9%)	1,356,480 (76.2%)
Surgical DRGs, N (%)	440,403 (24.4%)	17,359 (62.1%)	423,044 (23.8%)

*All results significant at $P < 0.0001$.

†Any PSI event includes the following PSIs: PSI #3 Pressure Ulcer, #5 Foreign Body Left in During Procedure, #6 Iatrogenic Pneumothorax, #7 Central Venous Catheter-related Bloodstream Infections, #8 Postoperative Hip Fracture, #9 Postoperative Hemorrhage or Hematoma, #10 Postoperative Physiologic and Metabolic Derangements, #11 Postoperative Respiratory Failure, #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis, #13 Postoperative Sepsis, #14 Postoperative Wound Dehiscence, and #15 Accidental Puncture or Laceration.

‡All-cause readmissions within 30 days from the index discharge.

DRGs indicates diagnosis-related groups; PSI, Patient Safety Indicator.

11.3%, respectively) (all comparisons were significant at $P < 0.0001$).

We explored differences in the top 10 reasons for readmission among index hospitalizations with and without any prior PSI event. Although the causes of readmission, as illustrated by CCS categories, were generally similar between the 2 groups (Table 3), there were some differences in prevalence rates for specific CCS categories. For example, "complications of surgical procedures or medical care" and "complication of device" were the top reasons for readmission (6.5% and 5.0% of readmissions, respectively) among hospitalizations with any prior PSI event; both of these occurred less frequently among those without any prior PSI event (4.8% and 2.8%, respectively). Similarly, although pneumonia and urinary tract infections were among the top reasons for readmission for hospitalizations with any prior PSI event (4.7% and 4.2%, respectively), these were not as frequent among those without any prior PSI event (3.7% and 2.4%, respectively) (all comparisons were significant at $P < 0.0001$). Moreover, although septicemia (another potentially hospital-related infection) occurred in 4.0% of readmissions associated with index hospitalizations that had any prior PSI event, it was not among the top 10 reasons for readmission in hospitalizations that did not have any prior PSI event.

The odds of readmission were 23% higher for index hospitalizations with any PSI event compared with those with no event (CI, 1.19–1.26) (Table 4). The odds of readmission were higher for all selected PSIs except Accidental Puncture or Laceration, ranging from 22% higher for Iatrogenic Pneumothorax (CI, 1.03–1.45) to 61% higher for Postoperative Wound Dehiscence (CI, 1.27–2.05). This suggests that index hospitalizations with a particular PSI, such as Postoperative Wound Dehiscence, had a greater likelihood of readmission than index hospitalizations that did not have that individual PSI (although they could have had other PSIs in the index hospitalization). For the composites, the odds of readmission ranged from 15% higher for the Technical Care composite (CI, 1.08–1.22), compared with those without any of the PSIs comprising this composite, to 37% higher for the Continuity of Care composite (CI, 1.26–1.50). Interestingly, among those with at least 1 PSI event, the odds of readmission were 1.15 (CI, 1.02–1.31) for those that had 2 or more PSI events. Nesting hospitalizations within patients, and including dummy variables for time, had no impact on results.

DISCUSSION

A few recent studies provide some empirical evidence that PSI events can increase the risk of readmission.[10,11] We explore this relationship in the VA health care system, which has a state-of-the-art electronic medical record system and a strong commitment to providing high-quality care and improving discharge planning.[32,33,34] Compared with other studies, our study population is broader, and we examine the types of PSIs associated with the greatest likelihood of readmission. Our study is also one of the first to ascertain whether experiencing a PSI event in the index hospitalization is associated with the risk of readmission.

TABLE 2. Readmission Rates for Index Hospitalizations With and Without Selected PSI Events (Fiscal Years 2003–2007)*

PSIs	Index Hospitalizations With Selected PSI Events (N)	30-Day Readmissions† (N)	%	Index Hospitalizations Without Selected PSI Events (N)	30-Day Readmissions (N)	%
PSI #6 Iatrogenic Pneumothorax	883	159	18.0	1,715,423	245,536	14.3
PSI #7 Central Venous Catheter-related Bloodstream Infections	2254	459	20.4	1,098,790	147,620	13.4
PSI #9 Postoperative Hemorrhage or Hematoma	1519	285	18.8	428,528	48,230	11.3
PSI #10 Postoperative Physiologic and Metabolic Derangements	593	141	23.8	257,062	29,183	11.4
PSI #11 Postoperative Respiratory Failure	2642	470	17.8	169,806	16,815	9.9
PSI #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis	4605	781	17.0	426,778	47,915	11.2
PSI #13 Postoperative Sepsis	916	176	19.2	84,355	10,999	13.0
PSI #14 Postoperative Wound Dehiscence	434	87	20.0	78,433	9023	11.5
PSI #15 Accidental Puncture or Laceration	5301	809	15.3	1,784,271	260,192	14.6
Technical Care composite‡	7297	1181	16.2	1,799,830	260,806	14.5
Continuity of Care composite§	3847	721	18.7	255,152	28,920	11.3
Mixed composite‖	7674	1384	18.0	1,780,120	255,695	14.4

*All results significant at $P < 0.0001$.

†All-cause readmissions within 30 days from the index discharge.

‡The Technical Care composite includes PSIs reflecting technical skill–based care: PSIs #5 Foreign Body Left in During Procedure, #9 Postoperative Hemorrhage or Hematoma, #14 Postoperative Wound Dehiscence, and #15 Accidental Puncture or Laceration.

§The Continuity of Care composite includes PSIs reflecting continuity of care and medical management of high-risk patients in the perioperative setting: PSIs #10 Postoperative Physiologic and Metabolic Derangements, #11 Postoperative Respiratory Failure, and #13 Postoperative Sepsis.

‖The Mixed composite includes a mix of surgical and medical PSIs: PSIs #6 Iatrogenic Pneumothorax and #7 Central Venous Catheter-related Bloodstream Infections, #8 Postoperative Hip Fracture, and #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis.

PSI indicates Patient Safety Indicator.

Our study has several important findings. First, our results are consistent with the literature suggesting the potential impact of safety-related events on increased risk of readmission and other outcomes.[4,8–11,22] Second, despite the relatively low prevalence of individual PSIs, index hospitalizations with selected PSIs had significantly higher readmission rates than those without PSIs. In some instances, the odds were almost double. Having multiple PSI events increased the odds even more. Readmission rates were also higher for index hospitalizations with PSI composites compared with those without the composites, and, similar to other studies, for index hospitalizations with any PSI event compared with those with no event.[10,11] The risk of readmission was about 60% higher for index hospitalizations in which a Postoperative Hemorrhage or Hematoma or Postoperative Wound Dehiscence occurred compared with those in which these particular PSIs did not occur. In contrast, Friedman et al[11] found the highest risk of 30-day readmission among patients who had a Postoperative Pulmonary Embolism/Deep Vein Thrombosis or Accidental Puncture or Laceration in the index admission, whereas Encinosa and Hellinger[10] found the highest risk of 90-day readmission among patients with PSIs related to infections (PSIs #7 and #13), Postoperative Physiologic and Metabolic Derangements, and Postoperative Respiratory Failure. Differences in study methods, patient populations, readmission definitions, and random variation may explain the impact of different PSIs in each study.

Third, we found that the occurrence of any PSI event in the index hospitalization was reflected in the reason for readmission. In general, index hospitalizations with any prior PSI event had a greater likelihood of readmission for complications related to surgical or medical care or implanted devices, or for acute problems such as infections that might be hospital acquired, compared with those without any prior PSI event. It is noteworthy that index hospitalizations without any prior PSI event were likely to be readmitted for exacerbations of chronic conditions, such as HF.

Our findings suggest that 1 way to reduce readmissions is to target and minimize potentially preventable AEs that occur in hospitals. The establishment of surgical checklists, training of staff in patient safety, hospital-wide efforts to reduce hospital-acquired infections, and developing incentives to improve care are just a few examples of possible steps hospitals can take to reduce readmissions.[35] Moreover, early follow-up and monitoring of surgical discharges may help prevent continuation of in-hospital complications into the postdischarge period and detect delayed complications early.[13]

Index hospitalizations with PSIs reflecting continuity of care had a higher likelihood of being readmitted than hospitalizations that did not have these PSIs. This composite comprises the same PSIs that had the greatest impact on likelihood of readmission in Encinosa and Hellinger's[10] study, providing further empirical support for the validity of this composite construction. As readmissions are sometimes thought of as "missed opportunities to better coordinate care,"[19] our results provide additional evidence that interventions to improve coordination of care across inpatient

TABLE 3. Top 10 Reasons for Readmission Among Index Hospitalizations With and Without Any Prior PSI Event*

	Readmissions After Index Hospitalizations With Any PSI Event (n = 6231)			Readmissions After Index Hospitalizations Without Any PSI Event (n = 255,791)		
	CCS (+)	Name	N (%)	CCS (+)	Name	N (%)
1	238	Complications of surgical procedures or medical care	402 (6.5)	108	Congestive heart failure; nonhypertensive	15,433 (6.0)
2	237	Complication of device; implant or graft	311 (5.0)	101	Coronary atherosclerosis and other heart disease	12,890 (5.0)
3	122	Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	295 (4.7)	238	Complications of surgical procedures or medical care	12,347 (4.8)
4	159	Urinary tract infections	259 (4.2)	122	Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	9462 (3.7)
5	2	Septicemia (except in labor)	246 (4.0)	127	Chronic obstructive pulmonary disease and bronchiectasis	8330 (3.3)
6	101	Coronary atherosclerosis and other heart disease	226 (3.6)	45	Maintenance chemotherapy; radiotherapy	7677 (3.0)
7	108	Congestive heart failure; nonhypertensive	212 (3.4)	237	Complication of device; implant or graft	7102 (2.8)
8	50	Diabetes mellitus with complications	161 (2.6)	106	Cardiac dysrhythmias	6806 (2.7)
9	114	Peripheral and visceral atherosclerosis	137 (2.2)	159	Urinary tract infections	6040 (2.4)
10	129	Aspiration pneumonitis; food/vomitus	132 (2.1)	55	Fluid and electrolyte disorders	6037 (2.4)

*Any PSI event includes PSIs #3 Pressure Ulcer, #5 Foreign Body Left in During Procedure, #6 Iatrogenic Pneumothorax, #7 Central Venous Catheter-related Bloodstream Infections, #8 Postoperative Hip Fracture, #9 Postoperative Hemorrhage or Hematoma, #10 Postoperative Physiologic and Metabolic Derangements, #11 Postoperative Respiratory Failure, #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis, #13 Postoperative Sepsis, #14 Postoperative Wound Dehiscence, and #15 Accidental Puncture or Laceration.

(+) CCS indicates Agency for Healthcare Research and Quality's Clinical Classification Software; PSI, Patient Safety Indicator.

TABLE 4. Logistic Regression Models for Any PSI Event, Selected PSIs, and PSI Composites Predicting Likelihood of 30-Day Readmission* (Fiscal Years 2003–2007)

	Odds Ratio	95% CI
Any PSI event†	1.23	1.19–1.26
PSI #6 Iatrogenic Pneumothorax	1.22	1.03–1.45
PSI #7 Central Venous Catheter-related Bloodstream Infections	1.53	1.38–1.70
PSI #9 Postoperative Hemorrhage or Hematoma	1.60	1.40–1.83
PSI #10 Postoperative Physiologic and Metabolic Derangements	1.53	1.26–1.86
PSI #11 Postoperative Respiratory Failure	1.39	1.25–1.54
PSI #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis	1.33	1.23–1.44
PSI #13 Postoperative Sepsis	1.32	1.12–1.57
PSI #14 Postoperative Wound Dehiscence	1.61	1.27–2.05
PSI #15 Accidental Puncture or Laceration	1.07	0.99–1.15
Technical Care composite‡	1.15	1.08–1.22
Continuity of Care composite§	1.37	1.26–1.50
Mixed composite‖	1.22	1.15–1.30

*Controlling for age, sex, the occurrence of any other PSI(s), and comorbidities [AHRQ comorbidity software (AHRQ website, version 3.5)].

†Any PSI event includes PSIs #3 Pressure Ulcer, #5 Foreign Body Left in During Procedure, #6 Iatrogenic Pneumothorax, #7 Central Venous Catheter-related Bloodstream Infections, #8 Postoperative Hip Fracture, #9 Postoperative Hemorrhage or Hematoma, #10 Postoperative Physiologic and Metabolic Derangements, #11 Postoperative Respiratory Failure, #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis, #13 Postoperative Sepsis, #14 Postoperative Wound Dehiscence, and #15 Accidental Puncture or Laceration.

‡The Technical Care composite includes PSIs reflecting technical skill–based care: PSIs #5 Foreign Body Left in During Procedure, #9 Postoperative Hemorrhage or Hematoma, #14 Postoperative Wound Dehiscence, and #15 Accidental Puncture or Laceration.

§The Continuity of Care composite includes PSIs reflecting continuity of care and medical management of high-risk patients in the perioperative setting: PSIs #10 Postoperative Physiologic and Metabolic Derangements, #11 Postoperative Respiratory Failure, and #13 Postoperative Sepsis.

‖The Mixed composite includes a mix of surgical and medical PSIs: PSIs #6 Iatrogenic Pneumothorax and #7 Central Venous Catheter-related Bloodstream Infections, #8 Postoperative Hip Fracture, and #12 Postoperative Pulmonary Embolism or Deep Vein Thrombosis.

CI indicates confidence interval; PSI, Patient Safety Indicator.

settings and between the inpatient and outpatient settings are important in reducing readmission rates.[36–38] In addition, patients at high risk of readmission, such as those with PSIs reflecting continuity of care, may be a good subset to target during and after hospitalization.

Our study has several important strengths. Unlike previous cross-sectional studies in this area that examined surgical cases only,[10,11] we used 5 years of VA nationwide data, and included both medical and surgical discharges, providing an opportunity to examine PSI events and readmissions on a large population of patients. Use of composite measures helped ensure adequate statistical power to test the association between PSIs and readmissions. The consistency of our results with those of the private sector[10,11] provides increased evidence of the impact of the PSIs on readmissions.

There were also some limitations. We lacked clinical data for risk adjustment, thereby limiting the ability to ascertain patients' severity of illness. We lacked current data; however, PSI and readmission rates have remained relatively stable over time, suggesting that use of more recent data

would not have affected our results.[39–42] In addition, use of administrative data is always subject to coding inaccuracies. For example, some PSI events were likely to be present on admission, and some of the conditions that were incorporated as comorbidities could have been complications of care.[26,43] However, several recent studies show that a number of the PSIs (including those with the highest impact on readmissions in our study) have moderate to good positive predictive ability based on chart review.[44–47] In addition, we did not examine organizational-level factors potentially associated with readmissions, because our primary focus was on the relationship between PSIs and readmissions. Future studies should examine their impact on readmissions. Finally, we could not establish causality between PSIs and readmissions because unobserved patient or provider factors could affect both measures, nor could we determine whether readmissions were clinically related to the index hospitalization, planned, or potentially preventable.

The relative lack of direct financial incentives in the VA compared with the private sector may result in less pressure to decrease readmission rates in the VA. However, the VA operates in an increasingly budget-constrained environment, and the public reporting of VA readmission rates should increase pressure to better understand the causes of readmissions and to develop interventions to reduce them. Although previous studies have linked readmissions with the quality of inpatient care during the index hospitalization,[48,49] our study suggests that focusing on AEs that occur during the index hospitalization provides another important way of viewing readmission as an indicator of quality of care. Thus, efforts to reduce readmissions should also incorporate quality improvement initiatives targeted at minimizing and preventing inpatient AEs.

REFERENCES

1. Institute of Medicine. *To Err is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 2002.
2. Agency for Healthcare Research and Quality. 2009 National Healthcare Quality Report. Available at: http://www.ahrq.gov/qual/nhqr09/nhqr09.pdf. Accessed November 3, 2011.
3. Miller MR, Elixhauser A, Zhan C, et al. Patient Safety Indicators: using administrative data to identify potential patient safety concerns. *Health Serv Res*. 2001;36:110–132.
4. Romano PS, Geppert JJ, Davies S, et al. A national profile of patient safety in U.S. hospitals. *Health Aff (Project Hope)*. 2003;22:154–166.
5. Weiner BJ, Alexander JA, Baker LC, et al. Quality improvement implementation and hospital performance on patient safety indicators. *Med Care Res Rev*. 2006;63:29–57.
6. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Qual Saf Health Care*. 2003;12(suppl 2):ii58–ii63.
7. Foster DA. Consequences and costs of medical injuries in Medicare inpatients. Paper presented at Academy Health annual research meeting. Seattle, WA. 2006.
8. Rivard P, Luther S, Christiansen C, et al. Using patient safety indicators to estimate the impact of potential adverse events on outcomes. *Med Care Res Rev*. 2008;65:67–87.
9. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA*. 2003;290:1868–1874.
10. Encinosa WE, Hellinger FJ. The impact of medical errors on ninety-day costs and outcomes: an examination of surgical patients. *Health Serv Res*. 2008;43:2067–2085.
11. Friedman B, Encinosa W, Jiang HJ, et al. Do patient safety events increase readmissions? *Med Care*. 2009;47:583–590.
12. Shufelt JL, Hannan EL, Gallagher BK. The postoperative hemorrhage and hematoma patient safety indicator and its risk factors. *Am J Med Qual*. 2005;20:210–218.
13. Weller WE, Gallagher BK, Cen L, et al. Readmissions for venous thromboembolism: expanding the definition of patient safety indicators. *Jt Comm J Qual Patient Saf*. 2004;30:497–504.
14. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med*. 2009;360:1418–1428.
15. Department of Veterans Affairs, Office of Assistant Deputy Under Secretary for Health. Survey of Veteran enrollees' health and reliance upon VA: 2011. Available at: http://www.va.gov/HEALTHPOLICYPLANNING/SOE2011/SoE2011_Report.pdf. Accessed November 14, 2011.
16. Centers for Medicare and Medicaid Services (CMS). CMS Hospital Compare Website. Available at: http://www.cms.gov/HospitalQualityInits/11_HospitalCompare.asp. Accessed November 14, 2011.
17. Department of Veterans Affairs. VA Hospital Compare. Available at: http://www.hospitalcompare.va.gov/and http://www.app.hospitalcompare.va.gov/. Accessed April 2012.
18. PriceWaterhouse Coopers' Health Research Institute. The Price of Excess: Identifying Waste in Healthcare. 2008. Available at: http://www.pwc.com/cz/en/verejna-sprava-zdravotnictvi/prices-of-excess-health care-spending.pdf. Accessed April 2012.
19. Payment policy for inpatient readmissions. *Report to Congress: promoting greater efficiency in Medicare*. Washington, DC: Medicare Payment Advisory Commission; 2007:103–115.
20. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011;306:1688–1698.
21. VA Information Resource Center. VIReC Research User Guide: FY2002 VHA Medical SAS Inpatient Datasets. Available at: http://www.virec.research.va.gov/References/RUG/RUG-Inpatient02.pdf. Accessed May 2012.
22. Rosen AK, Rivard P, Zhao S, et al. Evaluating the patient safety indicators: how well do they perform on Veterans Health Administration data? *Med Care*. 2005;43:873–884.
23. Bernheim SM, Lin Z, Grady JN, et al. *Measures Maintenance Technical Report: Acute Myocardial Infarction, Heart Failure, and Pneumonia 30-Day Risk-Standardized Readmission Measures*. New Haven, CT: Yale–New Haven Health Services Corporation/Center for Outcomes Research and Evaluation; 2011.
24. Agency for Healthcare Research and Quality. *AHRQ Quality Indicators: Guide to Patient Safety Indicators*. Rockville, MD: US Department of Health and Human Services, Agency for Healthcare Research and Quality; 2003.
25. Agency for Healthcare Research and Quality. PSI SAS Syntax software Version 4.1a,. Available at: http://www.qualityindicators.ahrq.gov/Archive/Software.aspx. Accessed October 20, 2011.
26. Houchens R, Elixhauser A, Romano P. How often are potential patient safety events present on admission? *Jt Comm J Qual Patient Saf*. 2008;34:154–163.
27. Chen Q, Rosen AK, Cevasco M, et al. Detecting patient safety indicators: how valid is "foreign body left during procedure" in the Veterans Health Administration? *J Am Coll Surg*. 2011;212:977–983.
28. Rosen A, Loveland S, Romano P, et al. Effects of resident duty hour reform on surgical and procedural patient safety indicators among hospitalized Veterans Health Administration and Medicare patients. *Med Care*. 2009;47:723–731.
29. Rivard P, Elwy R, Loveland S, et al. *Applying Patient Safety Indicators (PSIs) across health care systems: achieving data comparability. Vol 2*. Rockland, MD: Agency for Healthcare Research and Quality; 2005:7–25.
30. HCUP CCS Fact Sheet. Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality, Rockville, MD. January 2012. Available at: www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp. Accessed April 2012.
31. HCUP Comorbidity Software. Healthcare Cost and Utilization Project (HCUP). October 2011. Agency for Healthcare Research and Quality. Available at: http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp. Accessed October 20, 2011.
32. Asch SM, McGlynn EA, Hogan MM, et al. Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. *Ann Intern Med*. 2004;141:938–945.

33. Jha AK, Perlin JB, Kizer KW, et al. Effect of the transformation of the Veterans Affairs Health Care System on the quality of care. *N Engl J Med*. 2003;348:2218–2227.
34. Department of Veterans Affairs. Quality Enhancement Research Initiative Chronic Heart Failure VA Hospital To Home (H2H) Initiative. 2011.
35. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med*. 2009;360:491–499.
36. Coleman EA, Parry C, Chalmers S, et al. The care transitions intervention: results of a randomized controlled trial. *Arch Intern Med*. 2006;166:1822–1828.
37. Jack BW, Chetty VK, Anthony D, et al. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. *Ann Intern Med*. 2009;150:178–187.
38. Naylor MD, Brooten DA, Campbell RL, et al. Transitional care of older adults hospitalized with heart failure: a randomized, controlled trial. *J Am Geriatr Soc*. 2004;52:675–684.
39. Rosen AK, Shin MH, Loveland S, et al. Surgical Readmission Rates in the VA. Academy Health Annual Research Meeting, Boston, MA. June 2010.
40. Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: lessons from the Veterans Administration. *Med Care*. 2006;44:850–861.
41. Shin MH, Borzecki A, Loveland S, et al. Trends in Readmission Rates within the Veterans Health Administration. Poster Presentation. AcademyHealth Annual Research Meeting, Boston, MA. June 2010.
42. Raleigh VS, Cooper J, Bremner SA, et al. Patient safety indicators for England from hospital administrative data: case-control analysis and comparison with US data. *BMJ*. 2008;337–342.
43. Iezzoni LI. *Risk Adjustment for Measuring Health Care Outcomes*. Chicago: Health Administration Press; 2003.
44. Borzecki A, Kaafarani H, Cevasco M, et al. How valid is the AHRQ patient safety indicator postoperative hemorrhage or hematoma? *J Am Coll Surg*. 2011;212:946–953.
45. Cevasco M, Borzecki A, McCluskey D, et al. Positive predictive value of the AHRQ patient safety indicator postoperative wound dehiscence. *J Am Coll Surg*. 2011;212:962–967.
46. Utter GH, Cuny J, Sama P, et al. Detection of postoperative respiratory failure: how predictive is the Agency for Healthcare Research and Quality's Patient Safety Indicator? *J Am Coll Surg*. 2010;211:347–354. e341–329.
47. Rosen AK, Itani KM, Cevasco M, et al. Validating the patient safety indicators in the Veterans Health Administration: do they accurately identify true safety events? *Med Care*. 2012;50:74–85.
48. Ashton CM, Del Junco DJ, Souchek J, et al. The association between the quality of inpatient care and early readmission: a meta-analysis of the evidence. *Med Care*. 1997;35:1044–1059.
49. Weissman JS, Ayanian JZ, Chasan-Taber S, et al. Hospital readmissions and quality of care. *Med Care*. 1999;37:490–501.

Am J Cardiovasc Dis 2012;2(3):184-191
www.AJCD.us /ISSN:2160-200X/AJCD1205003

Original Article

Choices in the use of ICD-9 codes to identify stroke risk factors can affect the apparent population-level risk factor prevalence and distribution of CHADS2 scores

James A Rothendler[1,2], Adam J Rose[1,3], Joel I Reisman[1], Dan R Berlowitz[1,2,3], Lewis E Kazis[1,2]

[1]Center for Health Quality, Outcomes and Economic Research, Health Services Research and Development, Bedford VA Medical Center, Bedford, MA, USA; [2]Department of Health Policy and Management, Boston University School of Public Health, Boston, MA, USA; [3]Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA

Received May 18, 2012; Accepted June 10, 2012; Epub July 25, 2012; Published August 15, 2012

Abstract: While developed for managing individuals with atrial fibrillation, risk stratification schemes for stroke, such as CHADS2, may be useful in population-based studies, including those assessing process of care. We investigated how certain decisions in identifying diagnoses from administrative data affect the apparent prevalence of CHADS2-associated diagnoses and distribution of scores. Two sets of ICD-9 codes (more restrictive/ more inclusive) were defined for each CHADS2-associated diagnosis. For stroke/transient ischemic attack (TIA), the more restrictive set was applied to only inpatient data. We varied the number of years (1-3) in searching for relevant codes, and, except for stroke/TIA, the number of instances (1 vs. 2) that diagnoses were required to appear. The impact of choices on apparent disease prevalence varied by type of choice and condition, but was often substantial. Choices resulting in substantial changes in prevalence also tended to be associated with more substantial effects on the distribution of CHADS2 scores.

Keywords: Stroke, atrial fibrillation, risk stratification, CHADS2, ICD-9-CM codes

Introduction

Atrial fibrillation (AF) is a risk factor for ischemic stroke, but among patients with AF, the level of risk for an individual patient varies according to other clinical characteristics [1-5]. Several risk stratification schemes have been developed for assessing the relative risk of stroke in patients with non-valvular AF [2-8], and such information may help clinicians in making more informed decisions regarding clinical management. While these risk stratification schemes were developed for use in individual patients, they may also be useful in population-based studies [9-14]. For individual patients, chart reviews to obtain prior and current medical conditions are typically used for identifying characteristics associated with increased risks of stroke. However, for assessing relative risks for large cohorts of patients, administrative data is often more feasible.

There is disagreement on which risk stratification scheme provides the most useful information in patients with AF. We have chosen to focus on the CHADS2 because its elements (age and diagnoses) are often available in large administrative databases [6]. (Some risk stratification schemes use data that is less commonly available in such databases, such as systolic blood pressure or heart size measurements from echocardiograms [3-5].) In the CHADS2 scheme, a score is created by adding one point for each of the elements of Congestive heart failure (CHF), Hypertension, Age ≥75 or Diabetes, and two points for prior Stroke or transient ischemic attack (TIA) [6]. In the initial validation of CHADS2 in Medicare beneficiaries, a chart review was used to identify the presence of hypertension, a history of stroke or TIA and, for most of the cohort, diabetes [6]. However, International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9) codes were

used to identify "recent" CHF exacerbations associated with hospitalizations. ICD-9 codes were also used to identify the presence of diabetes for some of the patients. In a subsequent study comparing the CHADS2 to other risk adjustment schemes, any history of CHF was substituted for recent exacerbations [8].

Since the initial validation of CHADS2, a number of population-based studies have defined the elements of the CHADS2, in whole or in part, by the use of ICD-9 diagnosis codes [9-12]. However, in defining the various elements of the CHADS2, there have been differences among studies in the specifications of sets of ICD-9 codes associated with each condition, the number of times that a code was required to be present or the use of inpatient vs. outpatient datasets from which codes are extracted. Furthermore, there has been variation in the time interval used in searching for such codes.

Studies that have focused on validating sets of ICD-9 codes to represent the types of clinical conditions associated with the CHADS2 scheme have reported differences in sensitivity and specificity based on similar types of choices [15-19]. Use of ICD-9 codes to identify stroke has been particularly difficult in regard to identifying code sets that have both high sensitivity and specificity [19-21]. Therefore, among investigators who have used ICD-9-based datasets in defining CHADS2 scores, variation in codes to identify CHADS2 conditions is probably not surprising.

Differences in choices regarding how ICD-9 codes are used to identify the clinical conditions associated with CHADS2 could also impact the distribution of CHADS2 scores at the population level. This, in turn, could affect statistical associations and conclusions. The purpose of this study was to assess how different sets of ICD-9 codes as well as other decisions (such as the number of times a code appears and the time over which the presence of codes is searched) affect the apparent prevalence of CHADS2-associated conditions and the distribution of CHADS2 scores in a population of patients with non-valvular AF.

Methods

Datasets

Department of Veterans Affairs (VA) national datasets encompassing outpatient visits and/or acute inpatient hospitalizations were used to determine the presence of certain ICD-9 codes, as further specified below.

Population

Patients with "non-valvular" AF were identified from among those cared for in the VA during fiscal year (FY) 2007 (October 2006 through September 2007). Diagnoses for AF were based on having two or more inpatient or outpatient ICD-9 codes of 427.31 separated by at least 60 days. Patients were excluded if there was a history of valvular heart disease manifested by mitral stenosis or certain types of prior cardiac valve surgery [22], as indicated by the following ICD-9 diagnostic or procedure codes: 394.0, 394.2, 396.0, 396.1, 396.8, V42.2, V43.3, 35.10 to 35.14 or 35.20 to 35.28. These exclusion diagnoses were based on outpatient or inpatient records in the three years prior to the initial diagnosis of AF identified in FY2007.

ICD-9 definitions for CHADS2 components

A number of reports in the literature that incorporate CHADS2 scores have defined the presence of CHF, hypertension, diabetes or stroke/TIA using sets of ICD-9 codes [9-12]. Other reports, not involved with CHADS2, have studied ICD-9 code sets for these conditions in comparison to chart reviews [15, 16, 18, 19]. "Core" sets of codes often include 428.x for CHF, 401.xx-405.xx for hypertension and 250.x for diabetes. Additional codes for each condition, typically associated with other comorbidities or complications, have been used by some investigators. For example, the set for CHF has included 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.03, 404.13, 404.91, 404.93, that for hypertension has included 437.2, and that for diabetes has included 357.2, 362.0 and 366.41 [10, 19]. While it did not seem likely that use of these additional codes would identify a substantial number of additional patients with the relevant condition, we investigated this issue by specifying "more restrictive" and "more inclusive" code sets, as indicated in **Table 1**.

Compared to the other CHADS2 conditions, the use of ICD-9 codes to identify patients with prior ischemic stroke or TIA is more problematic. Because the CHADS2 scheme includes any prior history of ischemic stroke or TIA, it becomes

Table 1. ICD-9 codes used to define more restrictive and more inclusive sets associated with the CHADS2 conditions. TIA: transient ischemic attack

CHADS2 condition	More restrictive set	More inclusive set
Congestive heart failure	428.x	428.x, 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.03, 404.13, 404.91, 404.93
Hypertension	401.x, 402.x, 403.x, 404.x, 405.x	401.x, 402.x, 403.x, 404.x, 405.x, 437.2
Diabetes	250.x	250.x, 357.2, 362.0, 366.41
Stroke and TIA	433.x1,* 434.x1,* 435.x*	433.x1, 434.x1, 435.x, 436, 438.x

*Assessed using only inpatient data

important to search for records indicating prior in addition to acute events. Validation studies in the literature have typically assessed the codes associated with acute strokes associated with hospitalizations, and even in that setting, it has been difficult to identify code sets that have both good sensitivity and specificity [20, 21]. Using one year of inpatient records, Birman-Deych reported use of an ICD-9 code set that achieved a sensitivity of 58% for detecting acute or prior strokes, but patients in the cohort were selected based on a hospitalization for AF [19]. In attempting to identify any history of an ischemic stroke or TIA in a population that includes individuals without hospitalizations within a specified span of data, it would seem important to include outpatient records. However, we are not aware of published studies that have reported on systematic validation of outpatient ICD-9 codes for either acute or prior strokes (or TIAs). While we included outpatient codes in our more inclusive code set, as indicated below, we recognize that the accuracy of such codes is unknown at this time.

For the more restrictive code set for stroke/TIA, we included ICD-9 codes 433.x1 and 434.x1 identified from acute inpatient records, as these codes have been found to have high specificity and positive predictive value [20]. Also included in the restrictive set is 435.x for TIA. We accepted these codes in any position in the list of discharge diagnoses.

For the more inclusive code set, we included codes that appeared in either acute inpatient or in outpatient records. In addition to 433.x1, 434.x1 and 435.x, we included codes 436 and 438.x. Investigators who have defined CHADS2 using ICD-9 codes have often included code 436 ("Acute, but ill-defined, cerebrovascular disease") in identifying stroke [9-12], although it had a positive predictive value of only 48% in VA inpatient datasets [20]. We included ICD-9 code 438.x ("Late effects of cerebrovascular disease") in an effort to detect previous strokes [11, 19]. One potential issue is whether, in the outpatient setting, prior strokes are being coded with 438.x or with one of the codes more typically associated with acute events. Yiannakoulias et al found that in certain outpatient facilities in the province of Alberta, Canada, ICD-9 codes 435.x, 436 and 438.x were the most common of those for cerebrovascular disease [23]. However, there were substantial differences among the facilities in which of these codes predominated. For example, at some facilities, 438.x predominated, while at others, it was 436. While the actual medical conditions were not determined by chart review, the authors suggested that such wide variation in coding was unlikely to be explainable by differences in disease prevalence.

There was a change in the definition of ICD-9 code 436 beginning in October 2004 that was intended to remove the words "stroke" and "cerebrovascular accident" (CVA) from its description and to re-index those terms to other codes [24, 25]. However, it is unclear how this has changed coding practices in either the inpatient or outpatient setting, and some reference documents have continued to include, for certain sub-entries, the terms "stroke" and "CVA" in association with code 436 [26, 27]. Therefore we counted its presence in periods after September 2004 as part of the more inclusive code set. V12.54 is a current code for indicating prior stroke or TIA without residual sequelae, but was not relevant to this study as it was not an official code until Oct. 1, 2007.

In searching for relevant ICD-9 codes for each of the CHADS2 conditions, we assessed the effect

Table 2. Prevalence (%) of Components of CHADS2 by ICD-9 set (restrictive/inclusive), number of diagnoses required and lookback period. TIA: transient ischemic attack

Component Variant	Lookback Period					
	1 year		2 years		3 years	
	2 diagnoses	1 diagnosis	2 diagnoses	1 diagnosis	2 diagnoses	1 diagnosis
Congestive Heart Failure						
More Restrictive	17.2	24.5	21.7	28.2	23.7	30.3
More Inclusive	17.3	24.7	21.9	28.4	23.9	30.5
Hypertension						
More Restrictive	61.6	76.9	73.7	81.6	77.0	83.8
More Inclusive	61.6	76.9	73.7	81.6	77.0	83.8
Diabetes						
More Restrictive	30.2	34.3	32.7	35.5	33.4	36.1
More Inclusive	30.4	34.4	32.8	35.6	33.4	36.2
Stroke and TIA						
More Restrictive	N/A	0.6	N/A	1.0	N/A	1.3
More Inclusive	N/A	9.6	N/A	11.7	N/A	13.1

of looking back 1, 2 and 3 years ("lookback" times) prior to the initial diagnoses of AF in FY2007. Also, for diagnoses involving CHF, hypertension and diabetes, we explored the effect of requiring only 1 vs. 2 diagnoses during each of the relevant "lookback" times.

Statistical methods

We used descriptive analyses to indicate the prevalence of disease as a percent of all individuals in the cohort. In addition, we assessed the distribution of the CHADS2 scores (0, 1, 2, or 3+) as a percent of all scores in the cohort for each of the choices related to ICD-9 code sets, number of diagnoses and years of lookback.

Results

A total of 126,167 individuals met the study criteria for non-valvular AF. Their mean age was 74.0, and 98.4% were male. **Table 2** shows the prevalence of the individual conditions that comprise the CHADS2 as a function of the ICD-9 code set (more restrictive vs. more inclusive), the number diagnoses required (1 vs. 2) and the number of "lookback" years used in assessing for the presence of diagnoses (1, 2 or 3).

For CHF, hypertension and diabetes, the additional codes in the more inclusive set made relatively little difference, as expected, in the calculated disease prevalence. However, the requirement of having at least two diagnoses substantially decreased the apparent prevalence for all three conditions, especially with only a one year lookback, although the effect was greater for CHF and hypertension in comparison to diabetes. For Stroke/TIA, there was a substantial difference in the calculated prevalence of disease between the more restrictive and more inclusive code sets, mainly due to the inclusion in the more inclusive code set of diagnoses drawn from outpatient datasets. A majority of such outpatient codes were for 436, and there was also a substantial prevalence of 438.x and 435.x, whereas there were relatively few 433.x1 and 434.x1.

For all four of the condition categories, the number of years of lookback also had a substantial effect on the calculated disease prevalence. On a relative basis, the difference was most pronounced for stroke/TIA. Diabetes showed the least change as a function of years of lookback, perhaps because diabetes is so often a focus of office visits.

Table 3 shows the distribution of CHADS2 scores as function of ICD-9 criteria, number of diagnoses required and years of database lookback. In this case, the more restrictive and more inclusive categories include the corresponding code sets for each of the four conditions. There were moderate differences in the CHADS2 distribution between the more restrictive and more inclusive categories, mainly due to the prevalence differences for stroke/TIA. The effect of the more restrictive/ more inclusive categories was greater for the higher CHADS2 scores.

Varying the number of required diagnoses and

Table 3. Distribution of CHADS2 (%) scores by ICD-9 set (restrictive/inclusive), number of diagnoses required, and lookback period

CHADS2 Variant	1-Year Lookback Value of CHADS2				2-Year Lookback Value of CHADS2				3-Year Lookback Value of CHADS2			
	0	1	2	3+	0	1	2	3+	0	1	2	3+
More restrictive												
2 diagnoses*	12.1	32.7	36.5	18.7	8.4	27.5	39.8	24.3	7.5	25.9	40.0	26.5
1 diagnosis	6.8	26.2	40.5	26.5	5.5	23.9	40.4	30.2	5.0	22.7	40.0	32.3
More Inclusive												
2 diagnoses*	11.4	30.0	33.7	24.9	7.8	25.0	35.9	31.3	7.0	23.3	35.5	34.2
1 diagnosis	6.4	24.2	37.1	32.3	5.2	21.7	36.2	36.8	4.6	20.5	35.5	39.4

*Except for stroke/transient ischemic attack, which uses 1 diagnosis

the number of years of database lookback both had substantial effects on the distribution of the CHADS2 scores. As expected, there was a decrease in those with scores of 3+ when two diagnoses were required. Similarly, the number of patients with CHADS2 scores of 0 was substantially higher with one year compared to three years of lookback, and there was a corresponding decrease in the percentage of those with scores of 3+ with one compared to three years of lookback. For the more restrictive code sets, use of two diagnoses over one year was associated with 12.1% of patients with CHADS2 score of 0, compared to 5% when one diagnosis was required over three years.

Discussion

Because the more inclusive sets of ICD-9 codes for hypertension, CHF and diabetes differed from the more restrictive sets only by inclusion of codes that indicate certain complications or comorbidities, it is not surprising that there were only small differences in prevalence between these sets. However, for these same conditions, the requirements based on the number of years of lookback and number of diagnoses required both had more substantial effects on disease prevalence. The extent of these differences may vary among different healthcare systems and their associated administrative datasets.

For stroke and TIA, there were also substantial differences in apparent disease prevalence as a function of the number of years of lookback. There were also large relative differences in disease prevalence between more restrictive and more inclusive code sets. For the more inclusive code set, the increased prevalence was mainly due to the inclusion of codes in the outpatient setting, including those more typically associated with acute events. However, it is possible that in many cases of "acute" codes, the intent was to indicate a prior stroke or TIA. To the extent that prior strokes or TIAs were being diagnosed, the relative contribution of such data to the overall prevalence of stroke/TIA might be expected to have a larger effect in VA datasets compared to those of some other healthcare systems. This hypothesis is based on a report that most patients dually eligible for VA care and Medicare receive their initial care for acute strokes in non-VA hospitals, and such acute events would not typically be included in the VA acute hospitalization files [28]. As previously noted, we are not aware of any validation studies in the outpatient setting for using ICD-9 codes to identify stroke/TIA, and chart reviews were not part of this study. Therefore, the relative accuracy of different outpatient codes to identify stroke/TIA is unknown, and further studies in this area are warranted. However, the apparent prevalence of stroke/TIA of 13.1% using three years of data is similar to other studies that have determined the prevalence of these conditions in patients with AF [29, 30].

Consistent with the effect on disease prevalence, choices studied in this report regarding identification of conditions from ICD-9 codes also had a substantial effect on the distribution of the CHADS2 scores. Reports in the literature that have assessed CHADS2 scores based on ICD-9 codes have varied in these parameters, and this may affect comparison among studies [9-12].

Even in our "more inclusive" set of codes for stroke/TIA, we did not include such codes as 433.x0 or 434.x0 since these codes have been found to have poor positive predictive value for

stroke, although they may indicate other manifestations of cerebrovascular disease [20]. We also did not assess variation of disease prevalence or CHADS2 distribution as a function of the position of a particular diagnosis (i.e. primary vs. anywhere in the record). Such an assessment would likely have resulted in further variation in the observed parameters. For those ICD-9 codes identified from outpatient records, we included all outpatient services (including laboratory, radiology, etc) and not just face-to-face clinical encounters. This may also have affected the relative prevalence of these codes.

There are several additional issues that may limit the generalizability of our findings. The VA population in our cohort was 98.4% male, and the VA population tends to have a higher prevalence of comorbidities than non-VA populations. Also, as previously noted, a substantial percentage of patients dually eligible for VA care and Medicare receive initial care for strokes outside the VA. Therefore, the presence of outpatient codes suggestive of stroke/TIA may have had a greater influence on apparent disease prevalence compared to other health care systems. In addition to care for stroke, many VA patients receive at least some routine outpatient care outside the VA, and the VA datasets may therefore not reflect some comorbidities not addressed at VA clinic encounters. Regarding outpatient diagnoses, the VA uses an electronic medical record (EMR) to document care. The EMR's characteristics and the methods by which providers choose ICD-9 codes using this EMR may result in different distributions of codes compared to those found in other healthcare systems. Because of the presence of these issues, the actual numbers in this report derived from VA datasets are likely to be different in other systems of care. Although specific numbers may vary, it seems likely that in other healthcare systems, decisions in identifying CHADS2 conditions using ICD-9 codes can similarly affect calculated disease prevalence and distribution of scores.

Our goal in this report was not to define "optimal" algorithms to identify the presence of the various conditions associated with CHADS2 from administrative data, but rather to demonstrate the variation that can occur in the observed prevalence of these conditions as a function of choices made in their identification. While the CHADS2 often is used on an "n" of 1 basis for the clinical care of patients who might be at increased risk of stroke, the intent of this paper is to show the effect of different choices regarding use of ICD-9 codes in deriving the CHADS2 scores for application in population-based studies and not for individual patients.

Researchers performing population-based studies based on ICD-9 codes identified in administrative datasets should be aware of how choices in specifying components of stroke risk stratification schemes, such as the CHADS2, affect the distribution of risk scores. Further studies would be useful to assess how such choices affect the accuracy of administrative diagnostic codes for different healthcare systems. In particular, validation studies are needed regarding the accuracy of outpatient ICD-9 codes to indicate prior strokes and TIAs.

Acknowledgements

This material is based upon work supported in part by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Health Services Research and Development; the Boston University School of Public Health and a contract from Boehringer Ingelheim, Pharmaceuticals Inc. Dr. Rose is supported by a Career Development Award from the Veterans Affairs Health Services Research and Development Service.

The views expressed in this report are solely those of the authors and do not necessarily represent the position or policy of the Department of Veterans Affairs or the United States Government, Boston University School of Public Health or Boehringer Ingelheim, Pharmaceuticals, Inc.

Address correspondence to: Dr. Lewis E Kazis, Pharmaceutical Effectiveness Program, Center for Health Quality, Outcomes and Economic Research (CHQOER), 200 Springs Rd. (152), Building 70, Bedford, MA 01730-1114, USA. Tel: 781-687-2860; E-mail: lek@bu.edu

References

[1] Stroke Risk in Atrial Fibrillation Working Group. Independent predictors of stroke in patients with atrial fibrillation: a systematic review. Neurology 2007; 69: 546-554.

[2] Hughes M and Lip GY. Guideline Development Group, National Clinical Guideline for Management of Atrial Fibrillation in Primary and Secondary Care, National Institute for Health and Clinical Excellence. Stroke and thromboembolism in atrial fibrillation: a systematic review

of stroke risk factors, risk stratification schema and cost effectiveness data. Thromb Haemost 2008; 99: 295-304.
[3] Hart RG and Pearce LA. Current status of stroke risk stratification in patients with atrial fibrillation. Stroke 2009; 40: 2607-2610.
[4] Stroke Risk in Atrial Fibrillation Working Group. Comparison of 12 risk stratification schemes to predict stroke in patients with nonvalvular atrial fibrillation. Stroke; a journal of cerebral circulation 2008; 39: 1901-1910.
[5] Fang MC, Go AS, Chang Y, Borowsky L, Pomernacki NK, Singer DE; ATRIA Study Group. Comparison of risk stratification schemes to predict thromboembolism in people with nonvalvular atrial fibrillation. J Am Coll Cardiol 2008; 51: 810-815.
[6] Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA 2001; 285: 2864-2870.
[7] Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 2010; 137: 263-272.
[8] Gage BF, van Walraven C, Pearce L, Hart RG, Koudstaal PJ, Boode BS, Petersen P. Selecting patients with atrial fibrillation for anticoagulation: stroke risk stratification in patients taking aspirin. Circulation 2004; 110: 2287-2292.
[9] Choudhry NK, Soumerai SB, Normand SL, Ross-Degnan D, Laupacis A, Anderson GM. Warfarin prescribing in atrial fibrillation: the impact of physician, patient, and hospital characteristics. Am J Med 2006; 119: 607-615.
[10] Lakshminarayan K, Solid CA, Collins AJ, Anderson DC, Herzog CA. Atrial fibrillation and stroke in the general medicare population: a 10-year perspective (1992 to 2002). Stroke 2006; 37: 1969-1974.
[11] Kaatz S, Douketis JD, Zhou H, Gage BF, White RH. Risk of stroke after surgery in patients with and without chronic atrial fibrillation. J Thromb Haemost 2010; 8: 884-890.
[12] Henriksson KM, Farahmand B, Johansson S, Asberg S, Terént A, Edvardsson N. Survival after stroke--the impact of CHADS2 score and atrial fibrillation. Int J Cardiol 2010; 141: 18-23.
[13] Rose AJ, Berlowitz DR, Ash AS, Ozonoff A, Hylek EM, Goldhaber-Fiebert JD. The business case for quality improvement: oral anticoagulation for atrial fibrillation. Circ Cardiovasc Qual Outcomes 2011; 4: 416-424.
[14] Sandhu RK, Bakal JA, Ezekowitz JA, McAlister FA. Risk stratification schemes, anticoagulation use and outcomes: the risk--treatment paradox in patients with newly diagnosed non-valvular atrial fibrillation. Heart 2011; 97: 2046-2050.
[15] Humphries KH, Rankin JM, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? J Clin Epidemiol 2000; 53: 343-349.
[16] Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR. Identifying hypertension-related comorbidities from administrative data: what's the optimal approach? Am J Med Qual 2004; 19: 201-206.
[17] Miller DR, Safford MM and Pogach LM. Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. Diabetes Care 2004; 27 Suppl 2: B10-21.
[18] Quan H, Khan N, Hemmelgarn BR, Tu K, Chen G, Campbell N, Hill MD, Ghali WA, McAlister FA; Hypertension Outcome and Surveillance Team of the Canadian Hypertension Education Programs. Validation of a case definition to define hypertension using administrative data. Hypertension 2009; 54: 1423-1428.
[19] Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care 2005; 43: 480-485.
[20] Reker DM, Hamilton BB, Duncan PW, Yeh SC, Rosen A. Stroke: who's counting what? J Rehabil Res Dev 2001; 38: 281-289.
[21] Benesch C, Witter DM Jr, Wilder AL, Duncan PW, Samsa GP, Matchar DB. Inaccuracy of the International Classification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. Neurology 1997; 49: 660-664.
[22] Go AS, Hylek EM, Borowsky LH, Phillips KA, Selby JV, Singer DE. Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study. Ann Intern Med 1999; 131: 927-934.
[23] Yiannakoulias N, Svenson LW, Hill MD, Schopflocher DP, James RC, Wielgosz AT, Noseworthy TW. Regional comparisons of inpatient and outpatient patterns of cerebrovascular disease diagnosis in the province of Alberta. Chronic Dis Can 2003; 24: 9-16.
[24] ICD-9-CM Coordination and Maintenance Committee Meeting, April 1- 2, 2004. Available from: http://www.cdc.gov/nchs/data/icd9/agendaapril04%20revised.pdf.
[25] International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM): ICD-9-CM Tabular List of Diseases (FY05); File DTAB05. 2004; Available from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD9-CM/2004.
[26] International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). ICD-9-CM Files via FTP, 2009. ICD-9-CM Index to Diseases (FY10)/Index to Diseases and Injuries. File: DINDEX10; Available from: ftp://

ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD9-CM/2009

[27] Centers for Medicare and Medicaid Services. Diagnosis and Procedure Codes: Abbreviated and Full Code Titles. Version 27 Full and Abbreviated Code Titles - Effective October 1, 2009. See file: V27LONG_SHORT_DX_110909.xls. Available from: https://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/codes.html

[28] Shen Y, Findley PA, Maney M, Pogach L, Crystal S, Rajan M, Findley TW. Department of Veterans Affairs-Medicare dual beneficiaries with stroke: where do they get care? J Rehabil Res Dev 2008; 45: 43-51.

[29] Rose AJ, Ozonoff A, Henault LE, Hylek EM. Warfarin for atrial fibrillation in community-based practise. J Thromb Haemost 2008; 6: 1647-1654.

[30] ACTIVE Writing Group of the ACTIVE Investigators. Connolly S, Pogue J, Hart R, Pfeffer M, Hohnloser S, Chrolavicius S and Yusuf S. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. Lancet 2006; 367: 1903-1912.

Evidence-based design: part of evidence-based medicine?

Richard Saitz

10.1136/eb-2012-100934

Editor, *Evidence-Based Medicine*

Correspondence to:
Dr Richard Saitz
Section of General Internal Medicine, Boston University & Boston Medical Center, Boston, Massachusetts, USA;
rsaitz@bmjgroup.com

► http://dx.doi.org/10.1136/ebmed-2012-100785

In this issue of *Evidence-Based Medicine*, Shepley and Watson push the envelope of evidence-based medicine and healthcare.[1] Their *Perspective* emphasises that health professionals and designers (eg, architects) should collaborate to create healthcare facilities, and that such collaboration should lead to better staff, patient and family outcomes. But is this just 'knowledge-based gibberish' as a writer has misguidedly lamented in a news article that describes the misuse of the term evidence-based?[2] Is this just another example of a field wanting to claim the lofty mantle of having evidence to support its actions? I do not think so.

Healthcare facility construction is costly, and effects on clinical outcomes could include important benefits and risks for many. The most obvious examples would be infections from poor ventilation or from construction that leads to practitioner behaviours that increase such risks (eg, no sinks near exam rooms). Shepley and Watson note psychological outcomes too. In the USA, the Joint Commission accredits healthcare facilities and maintains standards for them, including requirements for the physical facilities themselves.[3]

Shepley and Watson's *Perspective* and the importance of design for healthcare outcomes beg another question: Are facility standards evidence-based?

Shepley and Watson point out that there are papers published in the peer-reviewed literature on the topic, and that there are research-based standards, and they lament the lack of medical-designer collaborations. Their review stops after counting articles, describing authorship (few medical/design collaborations) and providing case examples. But what of study design? What level of evidence is desired and feasible for evidence-based (healthcare facility) design (EBD)? To what extent should such decisions be subjected to randomised trials, n=1 trials, and even systematic reviews? Surely these are possible for some relevant questions, though likely difficult or impossible for others.

It seems clear that the location and logistics of healthcare delivery can impact important health outcomes. EBD seems to be in its infancy, and will no doubt face serious challenges, but it also seems worthy of serious consideration as healthcare continues to strive to become more evidence-based.

Competing interests None.

References

1. Shepley MM, Watson A. Evidence-based design: medical and design researcher collaboration. *EBM* 2012;18:2–4.
2. Beam A. Knowledge-based gibberish. The Boston Globe. 2011. http://articles.boston.com/2011-07-08/lifestyle/29752513_1_evidence-verbal-tics-reality (accessed 8 Jul 2012).
3. Joint Commission Resources. http://www.jcrinc.com/Joint-Commission-Requirements/Hospitals/#EC (accessed 8 Jul 2012).

ADDICTION SCIENCE & CLINICAL PRACTICE

EDITORIAL Open Access

Screening and brief intervention (SBI): has it hit the tipping point?

Richard Saitz*

In 1961, Chafetz [1] reported the results of a randomized trial of brief advice by a psychiatrist to patients with alcoholism in the Massachusetts General Hospital (MGH) emergency department (ED); 42% of patients in the advice arm reported to an alcohol clinic versus only 1% in the control group. Fifty years later, in the Liberty Hotel, the same space as the former Charles Street Jail in Boston where there was a "drunk tank" and across from that same MGH ED, over 200 researchers and clinicians gathered to present over 100 abstracts and plenary sessions on screening and brief intervention (SBI). In those 50 years, thousands of patients participated in randomized trials of SBI; the US Institute of Medicine (in 1990) encouraged identification and intervention for unhealthy alcohol use for people across the spectrum from risky use through dependence; the World Health Organization validated assessment tools and showed SBI's effectiveness in primary care settings; and the US Substance Abuse and Mental Health Services Administration allocated substantial funding for SBI for both alcohol and other drugs across many general health settings.

Research presented at the September 2011 International Network on Brief Interventions for Alcohol and Other Drugs (INEBRIA) conference in Boston was from around the globe, covered alcohol and other drugs, crossed a variety of health settings and practitioners, and showed the sophistication that has been reached in the field. Research discussed when, where, and for whom SBI has or might not have efficacy, how to implement SBI programs, adaptations of SBI, costs and effectiveness, and many other topics. Nonetheless, despite the excitement, breadth, and sophistication, the fact remains that few patients eligible for SBI receive the service, and, as a result, opportunities to improve health and save health care costs are missed. Most people with alcohol and drug use disorders receive no treatment.

At least in some cases, SBI is ready for dissemination. And, with solid evidence available, health reforms promised in the US and elsewhere, and an international discussion taking serious shape regarding the integration of care for medical, mental health, and substance use conditions, SBI may have reached a tipping point for dissemination and implementation as well as for small- and large-scale studies of remaining efficacy and effectiveness questions.

The editors of ASCP invited INEBRIA attendees and presenters to submit their studies for peer review and possible publication. The initial core of this thematic series, "Screening and brief intervention for unhealthy alcohol and other drug use," is the result of that call for papers, a seed that we expect will grow as the SBI literature continues to become more robust. Papers will cover a range of topics from what the efficacy of SBI really is, to effectiveness in people with mental health and drug use conditions, to adolescent SBI, to health professional attitudes towards SBI, to SBI implementation research and even to state-of-the-art SBI research protocol design, among others. We hope this series begins a long and serious conversation.

Received: 6 June 2012 Accepted: 15 June 2012
Published: 28 August 2012

Reference

1. Chafetz ME: **A procedure for establishing therapeutic contact with the alcoholic.** *Q J Stud Alcohol* 1961, **22**:325–328.

doi:10.1186/1940-0640-7-14
Cite this article as: Saitz: Screening and brief intervention (SBI): has it hit the tipping point? *Addiction Science & Clinical Practice* 2012 **7**:14.

Correspondence: Richard Saitz rsaitz@bu.edu
Boston University Schools of Medicine and Public Health/Boston Medical Center, Boston, MA, USA

ADDICTION SCIENCE & CLINICAL PRACTICE

EDITORIAL Open Access

Science to improve care for people affected by unhealthy alcohol and other drug use

Richard Saitz* and Jeffrey Samet

Why *Addiction Science & Clinical Practice?*

The last 20 years have brought two major changes in the field of addiction. First, the US Institute of Medicine (IOM) encouraged recognition of a spectrum of alcohol and other drug use that affects health and is not limited to those with the highest severity [1]. Unhealthy use (the spectrum from use that risks consequences through addiction [2]) among those without addiction is much more common than addiction itself. The second major shift also began in the 1990s with an emphasis on addiction as a chronic illness [3,4] and on unhealthy use as a health condition.

Although drug and excessive alcohol use clearly have social and other environmental determinants and consequences, there is little doubt that health care should play a major role in addressing them. In the health-care sector, attention to unhealthy substance use cannot be limited to highly specialized care settings; most patients with these conditions appear in general health settings where such problems are all too often ignored. In 2006, the IOM urged improvements in the quality of care for substance use conditions [5]. Responding to that report is a major impetus for the establishment of *Addiction Science & Clinical Practice (ASCP)*. We see this focus as unique among journals that address addictions.

History of *ASCP*

Addiction Science & Clinical Practice was founded in 2002 by the National Institute on Drug Abuse (NIDA), an agency of the US National Institutes of Health and the largest funder of drug-abuse research in the world. In 2011, NIDA discontinued ownership of the journal and transferred editorial control to us. We hope to see all of their subscribers and authors continuing the dialogue they established so well. Under NIDA's stewardship, *ASCP* excelled at bringing researchers and clinicians together to better translate science into practice and back again–to blend the two by publishing reviews and roundtables. We aim to take the next steps in this thematic direction: to serve as a forum for clinically relevant research and perspectives that contribute to improving the quality of care (prevention and treatment) for people with unhealthy alcohol, tobacco, or other drug use and addictive behaviors across a spectrum of clinical settings–from emergency departments to prisons, from primary-care offices to the Internet, and from general hospitals to specialty care programs.

A modern publishing model for *ASCP*

To achieve these aims, we have moved firmly into the cutting edge of 21st-century publishing practices by teaming with BioMed Central. Although some feel tied to print publications, we realize that, in 2012, most scientific reading occurs online. We are interested in having a broad reach and large impact, and online open-access journals are best able to meet these goals. In this publishing model, authors pay the costs to publish their work, which funders of research recognize as legitimate and expected. For an initial period, we are grateful to have support from NIDA that allows us to defray those costs for authors.

With BioMed Central as publisher, *ASCP* continues to be listed in Medline, to have articles deposited in national and international repositories, and to make our content freely and permanently available around the globe at no charge. Articles will be published immediately upon acceptance and listed in PubMed shortly thereafter. Because they are widely indexed, articles in *ASCP* are easily and freely discoverable through Internet search engines. We seek to have an impact beyond the usual calculations of citations; we want our articles to benefit people receiving care.

Building blocks

We appreciate the contributions of our exceptional group of Associate Editors and world-class Editorial Board who represent a range of disciplines (e.g., psychiatry, internal

* Correspondence: rsaitz@bu.edu
Boston University Schools of Medicine and Public Health/Boston Medical Center, Boston, MA, USA

medicine, psychology), research approaches (e.g., clinical trials, health services, behavioral sciences), and countries and offer their expertise on a variety of substances (e.g., alcohol, drugs, tobacco).

We look forward to reviewing your quality work and appreciate in advance your willingness to peer review submissions that, in some dimension, will improve the quality of health care for individuals with unhealthy substance use. It is with great humility and honor that we accept the task of providing thoughtful, timely, and responsive reviews for all submissions to *ASCP* as it embarks on its next phase. We know all the articles we publish will benefit from the wisdom, creativity, and hard work of our international, multidisciplinary substance use research colleagues.

Acknowledgements
The Editors acknowledge the superb efforts of Donna Vaillancourt, Managing Editor, in the preparation of this editorial for publication and in the launch of ASCP.

Received: 9 February 2012 Accepted: 15 March 2012
Published: 15 March 2012

References

1. Committee on Treatment of Alcohol Problems: *Institute of Medicine: Broadening the Base of Treatment for Alcohol Problems* Washington: The National Academies Press; 1990.
2. Saitz R: **Unhealthy alcohol use.** *N Engl J Med* 2005, **352**:596-607.
3. McLellan AT, Lewis DC, O'Brien CP, *et al*: **Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation.** *JAMA* 2000, **284**:1689-1695.
4. O'Brien CP, McLellan AT: **Myths about the treatment of addiction.** *Lancet* 1996, **347**:237-240.
5. Committee on Crossing the Quality Chasm: *Adaptation to Mental Health and Addictive Disorders, Institute of Medicine: Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series* Washington: The National Academies Press; 2006.

doi:10.1186/1940-0640-7-1
Cite this article as: Saitz and Samet: **Science to improve care for people affected by unhealthy alcohol and other drug use.** *Addiction Science & Clinical Practice* 2012 **7**:1.

nature genetics

Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways

To newly identify loci for age at natural menopause, we carried out a meta-analysis of 22 genome-wide association studies (GWAS) in 38,968 women of European descent, with replication in up to 14,435 women. In addition to four known loci, we identified 13 loci newly associated with age at natural menopause (at $P < 5 \times 10^{-8}$). Candidate genes located at these newly associated loci include genes implicated in DNA repair (*EXO1*, *HELQ*, *UIMC1*, *FAM175A*, *FANCI*, *TLK1*, *POLG* and *PRIM1*) and immune function (*IL11*, *NLRP11* and *PRRC2A* (also known as *BAT2*)). Gene-set enrichment pathway analyses using the full GWAS data set identified exoDNase, NF-κB signaling and mitochondrial dysfunction as biological processes related to timing of menopause.

Menopause is the cessation of reproductive function of the human ovaries. This life stage is associated with one of the major hormonal changes of women, characterized by a decline in secretion of estrogen, progesterone and, to a lesser degree, testosterone. It influences a woman's well-being and is associated with several major age-related diseases including cardiovascular disease, breast cancer, osteoarthritis and osteoporosis[1]. Ovarian aging is reflected by the continuous decline of the primordial follicle pool, which is established during fetal life, subsequently leading to endocrine changes owing to loss of the negative feedback from ovarian hormones on the hypothalamic-pituitary axis. In addition to follicle loss, oocyte quality diminishes with increasing age, which is believed to be due to greater meiotic nondisjunction[2]. Oocyte quality may be controlled at the time germ cells are formed during fetal life, but it may also reflect accumulated damage during reproductive life and/or age-related changes in granulosa cell–oocyte communication[3]. Although both oocyte quantity and quality decline with increasing age, it is unclear whether they are controlled by the same mechanisms and whether they decline in parallel.

The average age at natural menopause in women of Northern European descent is 50–51 years (range 40–60 years)[4]. Heritability estimates from twin and family studies for age at natural menopause range from 44% to 65% (refs. 5–8). Thus far most genetic association studies regarding age at menopause have focused on candidate genes[9] from the estrogen pathway[10,11] or vascular components[12,13]. Recently, two GWAS have newly identified five loci associated with age at natural menopause on chromosomes 5, 6, 13, 19 and 20 (refs. 14,15). These loci, however, explained <1.5% of the phenotypic variation of age at natural menopause, suggesting that additional loci of small effect will probably be discovered in larger samples. Therefore, we conducted a two-stage GWAS of women of European ancestry, combining the women from the two previous GWAS[14,15] with new participants for a total of 38,968 women from 22 studies in the discovery stage, and 14,435 women from 21 studies in the replication stage.

RESULTS

In our discovery stage of 38,968 women with natural menopause aged 40–60 (**Supplementary Tables 1** and **2**), we identified 20 regions with SNPs meeting the genome-wide significance criterion $P < 5 \times 10^{-8}$ (**Fig. 1**). Four of these loci confirmed earlier reports of associations on chromosomes 5, 6, 19 and 20 (refs. 14,15; regions 5b, 6a, 19a and 20, respectively, in **Table 1**) and 16 loci were previously unidentified. We did not confirm one reported association on chromosome 13 (13q34, rs7333181, $P = 0.12$). The overall genomic inflation factor was 1.03 (**Fig. 1**, inset; SNP with lowest P value from each region, **Table 1**). There was no between-study effect heterogeneity across discovery studies ($P > 0.05/20 = 0.0025$) for the 20 SNP associations presented. Within the Framingham Heart Study group, we tested for differences in effect size for the 20 SNPs in retrospectively and prospectively collected menopause age, and found no significant differences (data not shown). The effect sizes ranged from 0.17 years (8.7 weeks) to nearly 1 year (50.5 weeks) per each copy of the minor allele. We computed the effect sizes for dichotomized age at natural menopause in women from the Women's Genome Health Study (WGHS). For early menopause, we compared women with age at menopause <45 ($N = 745$) to those with age at menopause >45. For late menopause, we compared women with age at menopause >54 ($N = 1,632$) to those with age at menopause <54. The estimated odds ratios for early menopause for the menopause-decreasing allele ranged from 1.01 to 2.03. The estimated odds ratios for late menopause for the menopause-decreasing allele ranged from 0.52 to 0.96 (**Supplementary Table 3**). The top SNPs in regions 2c, 5a and 19b were >400 kb but <1 Mb from the top SNP in another region on the same chromosome. The top SNP in each of these primary regions had low linkage disequilibrium (LD; $r^2 < 0.5$) with the top SNP in the nearby region. To determine whether these associations were independent, we carried out a conditional association analysis in the discovery study samples with the most significant SNP from each of the primary 17 regions included as covariates in the analysis. For regions 5a and 19b (rs890835 and rs12461110, respectively), the effect estimates in the conditional analysis were unchanged compared

A full list of authors and affiliations appears at the end of the paper.

Received 21 July 2011; accepted 2 December 2011; published online 22 January 2012; doi:10.1038/ng.1051

with in the discovery analysis (differences of 0.3% and 4%, respectively), and the *P* values were genome-wide significant. However, for region 2c, the effect size was ~12.5% lower in the conditional analysis than in the initial analysis, and the SNP *P* value was no longer genome-wide significant ($P = 9.8 \times 10^{-7}$; **Table 1**), suggesting that the association with rs7606918 is not independent of the rs1018348 region 2b association. We attempted replication only for the 19 SNPs that represented independent regions that reached genome-wide significance ($P < 5 \times 10^{-8}$), thus we did not pursue replication of rs7606918.

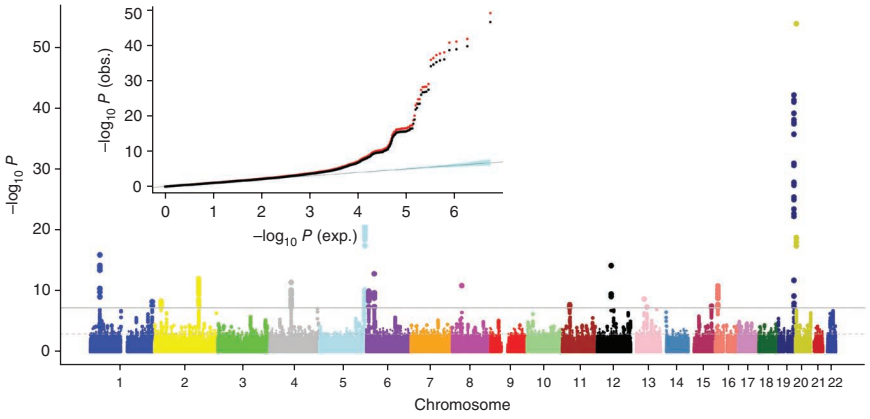


Figure 1 Discovery GWAS results. Manhattan plot of discovery meta-analysis. Inset, quantile-quantile plot of discovery primary analysis (red) and double genomic control–adjusted primary analysis (black). Obs., observed; exp., expected.

Replication

We used 21 studies contributing 14,435 women for replication of the 19 SNPs that defined the independent genome-wide significant regions from stage 1. We defined age at natural menopause using the same criteria as in the discovery studies (**Supplementary Table 1**). Of these studies, 17 ($n = 6,639$) were included in *in silico* replication (**Supplementary Table 2**); an additional 4 studies ($n = 7,796$) contributed *de novo* genotypes for the 19 SNPs (**Supplementary Table 2**; effect sizes and *P* values for replication and combined meta-analysis of discovery and replication samples, **Table 1**). There was no evidence for effect heterogeneity among the replication studies (**Table 1**). We also tested for heterogeneity between the *in silico* and *de novo* genotyped studies, and found no evidence for heterogeneity of effect (data not shown), suggesting that for the significant SNPs, the genotype imputation methods did not lead to significantly different effect size estimates than would have been obtained from direct genotyping. Of the 19 SNPs, 17 were genome-wide significant and had lower *P* values in combined meta-analysis of the discovery and replication samples. Regions 5a and 13a showed no evidence of association in the replication samples ($P > 0.50$) and were not genome-wide significant in combined discovery and replication meta-analysis. Four of the 17 replicated regions have been reported previously; thus our analysis identified 13 regions newly associated with age at natural menopause on the basis of genome-wide significant discovery with replication. In the combined discovery and replication meta-analyses, the effect estimates ranged from 8.2 to 49.3 weeks per minor allele. The estimated proportion of variance explained by the 17 replicated SNPs in the four replication studies with *de novo*–genotyped SNPs varied from 2.5% (Osteos) to 3.7% (EPOS and BWHHS) to 4.1% (PROSPECT-EPIC).

We used the largest study contributing data to our discovery GWAS (WGHS, $n = 11,379$) to explore whether substantial SNP-SNP interactions are present among the 17 replicated SNPs. We tested all 136 pairs of SNPs and found no evidence for interaction (all $P > 0.01$).

Roles of genes at or near newly identified loci

All but two of the replicated SNPs are intronic or exonic to known genes (**Table 2**). The top SNPs in regions 6b, 12, 19b and 20 are missense polymorphisms. Three of the four have been predicted to have damaging protein function by SIFT[16], and one by PolyPhen2 (ref. 17). Using dbSNP and LocusZoom[18], we identified the genes underlying the newly identified top regions. We used SCAN (see URLs) to identify all genes with SNPs that are in LD ($r^2 > 0.5$) with our SNPs (**Table 2**). We identified all SNPs with $r^2 \geq 0.8$ with our top SNPs and used several databases to determine whether the SNPs are associated with expression (**Table 2**).

The strongest new signal was on chromosome 4 (region 4, rs4693089; $P = 2.4 \times 10^{-19}$). The SNP is located in an intron of *HELQ*, which encodes the protein HEL308, a DNA-dependent ATPase and DNA helicase[19]. The second strongest new signal was on chromosome 12 (region 12, rs2277339; $P = 2.5 \times 10^{-19}$). This SNP is a nonsynonymous variant in exon 1 of *PRIM1*. The top SNP was significantly associated with expression of *PRIM1* in visual cortex, cerebellum and prefrontal cortex (**Table 2**).

Several other previously unidentified signals are located in introns of genes for which mouse models exist. These were region 8 in *ASH2L* (rs2517388; $P = 9.3 \times 10^{-15}$), region 15 in *POLG* (rs2307449; $P = 3.6 \times 10^{-13}$) and region 1b in *EXO1* (rs1635501; $P = 8.5 \times 10^{-10}$). *ASH2L* encodes a trithorax group protein, and is involved in X chromosome inactivation in women[20]. *POLG* encodes the catalytic subunit of mitochondrial DNA polymerase, the enzyme responsible for replication and repair[21] of mitochondrial DNA. *EXO1* is a member of the RAD2 nuclease family of proteins, which is involved in DNA replication, repair and recombination, and the top hit is in LD ($r^2 = 0.83$) with a functional polymorphism in *EXO1* that affects a transcription factor–binding site in the promoter. Region 11 (rs12294104; $P = 1.5 \times 10^{-11}$) is near and in LD ($r^2 = 0.92$) with SNPs in *FSHB*. Transcription of *FSHB* limits the rate of production of the heterodimeric follicle-stimulating hormone (FSH), a key pituitary gland–expressed hormone that stimulates maturation of follicles. Region 19a (rs11668344; $P = 1.5 \times 10^{-59}$) is in tight LD with SNPs in *IL11*; this cytokine stimulates the T cell–dependent development of immunoglobulin-producing B cells.

The top SNPs in two other previously unknown regions are nonsynonymous coding variants. Region 6b, rs1046089 ($P = 1.6 \times 10^{-16}$), is in exon 22 of *PRRC2A* and was associated with expression of several transcripts in the human leukocyte antigen (HLA) region in several tissues (**Table 2**). Region 19b, rs12461110 ($P = 8.7 \times 10^{-10}$) is in exon 5 of *NLRP11*. *PRRC2A* encodes HLA-B associated transcript 2 and has several microsatellite repeats. *NLRP11* encodes the nucleotide-binding domain and leucine-rich repeat–containing (NLR) family pyrin domain–containing 11 protein, which is implicated in the activation of proinflammatory caspases[22].

Of the remaining five new regions, the top SNPs for regions 1a, 2a, 2b and 13b are located in introns. These were rs4246511 in *RHBDL2* (0.24 years per minor allele, $P = 9.1 \times 10^{-17}$), which is thought to function as an intramembrane serine protease; rs2303369 in *FNDC4*, which encodes fibronectin type III domain–containing 4 ($P = 2.3 \times 10^{-12}$);

Table 1 Discovery and replication results

Region number	Most significant SNP	Chr.	Location (bp)	Minor/major	Analysis	MAF	Effect per minor allele (years)	s.e.m.	*P*	Absolute effect per minor allele (weeks)	Heterogeneity *P* value
1a	rs4246511	1	39152972	T/C	Discovery	0.271	0.289	0.035	1.02×10^{-16}	15.0	0.288
					Replication	0.293	0.133	0.052	0.01	6.9	0.680
					Combined		0.240	0.029	$\mathbf{9.08 \times 10^{-17}}$	12.5	0.012
1b	rs1635501	1	240107398	C/T	Discovery	0.478	–0.188	0.032	4.90×10^{-9}	9.8	0.080
					Replication	0.456	–0.110	0.048	0.023	5.7	0.016
					Combined		–0.164	0.027	$\mathbf{8.46 \times 10^{-10}}$	8.5	0.178
2a	rs2303369	2	27568920	T/C	Discovery	0.388	–0.174	0.030	3.80×10^{-9}	9.0	0.639
					Replication	0.391	–0.179	0.047	**0.000138**	9.3	0.542
					Combined		–0.175	0.025	$\mathbf{2.25 \times 10^{-12}}$	9.1	0.927
2b	rs10183486	2	171699217	T/C	Discovery	0.366	–0.219	0.031	7.88×10^{-13}	11.4	0.430
					Replication	0.362	–0.141	0.047	0.003	7.3	0.727
					Combined		–0.196	0.026	$\mathbf{2.21 \times 10^{-14}}$	10.2	0.169
2c[a]	rs7606918	2	172603695	G/A	Discovery	0.161	–0.228	0.041	2.89×10^{-8}	11.8	0.374
					Replication not attempted						
4	rs4693089	4	84592646	G/A	Discovery	0.486	0.209	0.030	3.28×10^{-12}	10.9	0.336
					Replication	0.492	0.273	0.047	$\mathbf{6.69 \times 10^{-9}}$	14.2	0.298
					Combined		0.228	0.025	$\mathbf{2.38 \times 10^{-19}}$	11.8	0.254
5a[b]	rs890835	5	175888877	A/C	Discovery	0.112	0.266	0.047	1.17×10^{-8}	13.8	0.003
					Replication	0.115	–0.037	0.072	0.613	1.9	0.486
					Combined		0.177	0.039	6.10×10^{-6}	9.2	0.0004
5b	rs365132	5	176311180	T/G	Discovery	0.490	0.275	0.029	1.90×10^{-21}	14.3	0.115
					Replication	0.494	0.319	0.046	$\mathbf{4.26 \times 10^{-12}}$	16.6	0.728
					Combined		0.287	0.025	$\mathbf{9.11 \times 10^{-32}}$	14.9	0.416
6a	rs2153157	6	11005474	A/G	Discovery	0.492	0.184	0.028	9.40×10^{-11}	9.5	0.858
					Replication	0.500	0.116	0.046	0.012	6.0	0.329
					Combined		0.165	0.024	$\mathbf{7.76 \times 10^{-12}}$	8.6	0.211
6b	rs1046089	6	31710946	A/G	Discovery	0.353	–0.226	0.031	1.31×10^{-13}	11.8	0.426
					Replication	0.358	–0.181	0.049	$\mathbf{1.91 \times 10^{-4}}$	9.4	0.732
					Combined		–0.213	0.026	$\mathbf{1.63 \times 10^{-16}}$	11.1	0.427
8	rs2517388	8	38096889	G/T	Discovery	0.174	0.274	0.040	1.13×10^{-11}	14.2	0.670
					Replication	0.189	0.234	0.062	$\mathbf{1.52 \times 10^{-4}}$	12.2	0.708
					Combined		0.262	0.034	$\mathbf{9.31 \times 10^{-15}}$	13.6	0.591
11	rs12294104	11	30339475	T/C	Discovery	0.172	0.226	0.040	1.63×10^{-8}	11.8	0.721
					Replication	0.180	0.223	0.060	$\mathbf{2.18 \times 10^{-4}}$	11.6	0.239
					Combined		0.225	0.033	$\mathbf{1.46 \times 10^{-11}}$	11.7	0.970
12	rs2277339	12	55432336	G/T	Discovery	0.102	–0.394	0.051	5.99×10^{-15}	20.5	0.088
					Replication	0.105	–0.347	0.077	$\mathbf{6.89 \times 10^{-6}}$	18.0	0.765
					Combined		–0.380	0.042	$\mathbf{2.47 \times 10^{-19}}$	19.7	0.610
13a	rs3736830	13	49204222	G/C	Discovery	0.157	–0.243	0.040	1.75×10^{-9}	12.6	0.859
					Replication	0.165	–0.033	0.062	0.594	1.7	0.905
					Combined		–0.180	0.034	9.41×10^{-8}	9.4	0.005
13b	rs4886238	13	60011740	A/G	Discovery	0.334	0.172	0.031	3.76×10^{-8}	8.9	0.974
					Replication	0.333	0.166	0.049	$\mathbf{6.09 \times 10^{-4}}$	8.6	0.953
					Combined		0.170	0.026	$\mathbf{9.53 \times 10^{-11}}$	8.9	0.919
15	rs2307449	15	87664932	G/T	Discovery	0.405	–0.167	0.030	2.59×10^{-8}	8.7	0.442
					Replication	0.387	–0.225	0.047	$\mathbf{1.61 \times 10^{-6}}$	11.7	0.328
					Combined		–0.184	0.025	$\mathbf{3.56 \times 10^{-13}}$	9.6	0.294
16	rs10852344	16	11924420	C/T	Discovery	0.415	0.198	0.029	1.28×10^{-11}	10.3	0.014
					Replication	0.426	0.093	0.046	0.042	4.8	0.599
					Combined		0.168	0.025	$\mathbf{1.01 \times 10^{-11}}$	8.7	0.054
19a	rs11668344	19	60525476	G/A	Discovery	0.363	–0.416	0.030	5.94×10^{-43}	21.6	0.112
					Replication	0.360	–0.415	0.048	$\mathbf{2.65 \times 10^{-18}}$	21.6	0.517
					Combined		–0.416	0.026	$\mathbf{1.45 \times 10^{-59}}$	21.6	0.987
19b[c]	rs12461110	19	61012475	A/G	Discovery	0.356	–0.174	0.030	9.49×10^{-9}	9.1	0.835
					Replication	0.344	–0.117	0.049	0.018	6.1	0.542
					Combined		–0.158	0.026	$\mathbf{8.74 \times 10^{-10}}$	8.2	0.320
20	rs16991615	20	5896227	A/G	Discovery	0.069	0.971	0.062	1.16×10^{-54}	50.5	0.356
					Replication	0.070	0.896	0.096	$\mathbf{7.90 \times 10^{-21}}$	46.6	0.088
					Combined		0.948	0.052	$\mathbf{1.42 \times 10^{-73}}$	49.3	0.509

Chr., chromosome. MAF, minor allele frequency. Heterogeneity, *P* values for heterogeneity among discovery studies and replication studies, and comparing all discovery to all replication studies. *P* value, replication *P* values that meet the criterion $P < 0.05/19 \approx 0.026$ are in bold. Combined analysis *P* values that reached genome-wide significance are in bold. [a]Conditional analysis: beta (SE): –0.199 (0.041); $P = 9.8 \times 10^{-7}$. [b]Conditional analysis beta (SE): 0.267 (0.046); $P = 6.5 \times 10^{-9}$. [c]Conditional analysis: beta (SE): –0.168 (0.031); $P = 3.8 \times 10^{-8}$.

Table 2 Characteristics of top SNP in each region

Region	SNP ID	Chr.	Location (bp)	Gene	Feature	Other ref. genes <60 kb from SNP	Genes with SNPs in LD (r^2 > 0.5) with SNP (gene symbol:r^2)	eQTL
1a	rs4246511	1	39152972	*RHBDL2*	Intron	*MYCBP*, *GJA10*, *RRAGC*	*MYCBP*:0.678; *RHBDL2*:0.678; *RRAGC*:0.678; GJA9:0.678; *LOC100130627*:0.678	
1b	rs1635501	1	240107398	*EXO1*	Intron		*LOC100131576*:0.678; *WDR64*:0.923; *LOC100133057*:1; EXO1:1	
2a	rs2303369	2	27568920	*FNDC4*	Intron	*GCKR*, *KRTCAP3*, *IFT172*, *NRBP1*	*FTHL3P*:0.967; GCKR:1; *GTF3C2*:0.967; *MPV17*:0.967; PPM1G:0.967; *UCN*:0.841; EIF2B4:0.967; *SNX17*:0.967; IFT172:1; *NRBP1*:0.967; TRIM54:0.841; *FNDC4*:1; C2orf16:0.513; *ZNF513*:0.967; KRTCAP3:1; *DNAJC5G*:0.841; *LOC100130981*:0.901	r^2 > 0.9 with multiple eSNPs for *IFT172* in lymph, adipose and blood; r^2 > 0.9 with multiple eSNPs for *KRTCAP3* in lymph and CD4; r^2 > 0.9 with eSNP for *SNX17* in PFC
2b	rs10183486	2	171699217	*TLK1*	Intron		*GORASP2*:0.929; TLK1:0.964	r^2 > 0.9 with eSNPs for *TLK1* in LCL and PFC
4	rs4693089	4	84592646	*HELQ*	Intron	*MRPS18C*, *FAM175A*	*MRPS18C*:1; *FAM175A*:1; *AGPAT9*:1; *HEL308*:1; *OK/SW-CL.36*:1	r^2 > 0.80 with eSNPs for *MRPS18C* in lymph and PFC; r^2 > 0.80 for *AGPAT9* in fibroblast
5a	rs890835	5	175888877	*RNF44*	Intron	*UBXD8*, *PCLKC*	*SNCB*:0.318; *RNF44*:1; *FAF2*:1; *PCDH24*:1	
5b	rs365132	5	176311180	*UIMC1*	Coding-synonymous	*HK3*	*FGFR4*:0.871; *HK3*:0.967; *ZNF346*:0.966; *UIMC1*:0.967; *UNC5A*:0.932	eSNP for *Hs.484258* in lymph; for *FGFR4* and *ZNF346* in PFC; for *ZNF346* in VC
6a	rs2153157	6	11005474	*SYCP2L*	Intron	*GCM2*	*MAK*:0.602; *GCM2*:0.602; *SYCP2L*:0.602	eSNP for *SYCP2L* in monocytes
6b	rs1046089[a]	6	31710946	*PRRC2A*	Missense	*BAT3*, *LST1*, *C6orf47*, *APOM*, *AIF1*, *NCR3*, *LY6G5C*, *LTB*, *BAT5*, *CSNK2B*, *LY6G5B*, *BAT4*, *TNF*	*AIF1*:0.963; *CLIC1*:0.682; *CSNK2B*:0.963; *HSPA1A*:0.649; *HSPA1B*:0.587; *HSPA1L*:0.649; *LTB*:0.963; *MSH5*:0.682; *NEU1*:0.587; *VARS*:0.682; *BAT2*:0.963; *BAT3*:0.963; *BAT4*:0.963; *BAT5*:0.963; *LST1*:0.963; *DDAH2*:0.682; *SNORD52*:0.587; *SNORD48*:0.587; *C6orf48*:0.587; *APOM*:0.963; *LSM2*:0.649; *C6orf47*:0.963; *LY6G5B*:0.963; *LY6G6D*:0.963; *LY6G6E*:0.963; *SLC44A4*:0.587; *C6orf27*:0.682; *C6orf25*:0.682; *LY6G6C*:0.961; *LY6G5C*:0.963; *NCR3*:0.963; *LY6G6F*:0.963; *C6orf26*:0.682; *SNORA38*:0.963; *LOC100130756*:0.963	eSNP for *LY6G5C* in CD4 and lymph; for *HLA-DRB4* in monocytes; for *C6orf10* in VC; for *AIF1* in lymph; for *HLA-DQA1* in LCL
8	rs2517388	8	38096889	*ASH2L*	Intron	*BAG4*, *EIF4EBP1*, *LSM1*, *STAR*	*STAR*:0.831; *ASH2L*:0.831; *LSM1*:0.831	
11	rs12294104	11	30339475	–		*MPPED2*, *C11orf46*	*FSHB*:0.92; *C11orf46*:1	r^2 > 0.9 with eSNP for *C11orf46* in lymph
12	rs2277339[b]	12	55432336	*PRIM1*	Missense	*HSD17B6*, *NACA*	–	eSNP for *PRIM1* in VC, CR and PFC
13a	rs3736830	13	49204222	*KPNA3*	Intron	*EBPL*	*KPNA3*:0.734; *EBPL*:0.734; *ARL11*:0.623; *LOC100131941*:0.623	
13b	rs4886238	13	60011740	*TDRD3*	Intron		*TDRD3*:1; *LOC390407*:0.731	r^2 > 0.8 with eSNP in adipose for *TDRD3*
15	rs2307449	15	87664932	*POLG*	Intron	*FANCI*	*POLG*:0.965; *RLBP1*:0.898; *ABHD2*:0.898; *FANCI*:0.965; *LOC728003*:0.898; *LOC100131654*:0.683	r^2 > 0.8 with eSNP for *RLBP1* in PFC
16	rs10852344	16	11924420	–		*TNFRSF17*, *RUNDC2A*, *GSPT1*	*TNFRSF17*:0.662; *GSPT1*:1; *COX6CP1*:1; *RSL1D1*:1; *ZC3H7A*:0.701; *RUNDC2A*:0.662; *LOC729978*:1	
19a	rs11668344	19	60525476	*TMEM150B*	Intron	*BRSK1*, *HSPBP1*, *COX6B2*, *LOC284417*, *IL11*, *SUV420H2*	*IL11*:0.962; *SAPS1*:0.894; *HSPBP1*:0.962; *BRSK1*:0.962; *SUV420H2*:0.962; *COX6B2*:0.962; *LOC284417*:0.962; *FAM71E2*:0.962	r^2 > 0.8 with eSNP for *MGC2705* in adipose and blood
19b	rs12461110[c]	19	61012475	*NLRP11*	Missense	*NLRP4*	*NLRP4*:0.514; *NLRP11*:0.514; *RFPL4A*:0.514; *LOC646663*:0.514; *LOC729974*:0.514	
20	rs16991615[d]	20	5896227	*MCM8*	Missense	*CRLS1*, *CHGB*, *TRMT6*	–	

CD4, $CD4^+$ lymphocyte cells; PFC, prefrontal cortex cells; LCL, lymphoblastoid cell line; CR, cerebellum; VC, visual cortex cells.
[a]Arg > His; predicted to be damaging (SIFT), benign (Polyphen2). [b]Asp > Ala; predicted to be damaging (SIFT), probably damaging (Polyphen2). [c]Pro > Leu; predicted to be damaging (SIFT), benign (Polyphen2). [d]Glu > Lys; predicted to be tolerated (SIFT), benign (Polyphen2).

rs10183486 in *TLK1* ($P = 2.2 \times 10^{-14}$), a nuclear serine-threonine kinase that is potentially involved in the regulation of chromatin assembly; and rs4886238 in *TDRD3* ($P = 9.5 \times 10^{-11}$). *TDRD3* is a binding partner for *FMR1*, which has been associated with primary ovarian insufficiency (POI). The top SNP in the final newly identified region, 16, is within 60 kb of three genes, *TNFRSF17*, *GSPT1* and *RUNDC2A*. It is in LD ($r^2 > 0.5$) with SNPs in these three genes and four others (rs10852344; $P = 1.0 \times 10^{-11}$; **Table 2**).

Pathway analyses

We used three independent pathway-based methods to identify connections among our single-marker associations and link them with broader biological processes. Although all three approaches (ingenuity pathway analysis (IPA, see URLs), MAGENTA[23] and GRAIL[24]) are based on published data, thus linking the gene products of our top hits to each other in functional pathways, each uses a substantially different methodology and uses different aspects of our results as input. Thus, we expect complementary results from the three approaches.

We used IPA (see URLs) to identify potential biological pathways common to the 17 replicated SNPs. On the basis of the genes physically nearest the 17 loci, we identified four major functional networks applying direct interactions only (**Supplementary Table 4**). Network 1, related to lipid metabolism, molecular transport and small molecule biochemistry, contained 14 of the genes nearest the menopause loci ($P = 1 \times 10^{-30}$). Central to this network is the *HNF4A* gene, which has a role in diabetes. Network 2, containing 12 of the input genes, relates to cell cycle, cell death and cancer ($P = 1 \times 10^{-24}$). The *ESR1* gene is central in this network, suggesting that genes in this network influence or are influenced by estrogen signaling. Network 3 is also partially related to cell death, and includes TNF and NF-κB ($P = 1 \times 10^{-19}$). Network 4 relates to infection mechanism, DNA replication, recombination and repair and gene expression ($P = 1 \times 10^{-12}$). Notably, several of the input genes included in network 1 (*EXO1* and *HELQ*) and network 2 (*UIMC1*, *FANCI* and *TLK1*) are also involved in DNA repair mechanisms.

We used gene set enrichment analysis (GSEA) implemented in MAGENTA[23] to explore pathway-based associations using the full GWAS results. Three pathways reached study-wide significance (false discovery rate (FDR) < 0.05), exoDNase ($P = 0.0005$), NF-κB signaling ($P = 0.0006$) and mitochondrial dysfunction ($P = 0.0001$; **Supplementary Table 5**).

Finally, we used the GRAIL method of literature-based pathway analysis[24] to explore the connections between genes near our top SNPs. Genes are considered related if they share informative words. GRAIL scores for genes associated with three of the replicating genome-wide significant SNPs were significant, *EXO1*, *FKBPL* and *BRSK1*. When we applied this method to a deeper set of 66 SNPs from the discovery meta-analysis with significance meeting FDR < 0.05, 12 genes had significant GRAIL scores: *EXO1*, *MSH6*, *PARL*, *RHBDL2*, *FKBPL*, *TP53BP1*, *TLK1*, *RAD54L*, *CHEK2*, *H2AFX*, *APEX1* and *REV3L*. *BRSK1* was also borderline significant with GRAIL FDR = 0.06.

Candidate genes

Within the discovery GWAS, 18,327 SNPs were within 60 kb of the start and end of transcription of 125 candidate genes selected because of a reported relationship with ovarian function (**Supplementary Table 6**). After multiple testing correction, 101 SNPs in or near five of the candidate genes (*DMC1*, *EIF2B4*, *FSHB*, *POLG* and *RFPL4A*) were significantly associated with age at natural menopause. SNPs in or near four of these genes were already identified as genome-wide significant (*EIF2B4*, region 2a; *RFPL4A*, region 19b; *POLG*, region 15; and *FSHB*, region 11). For the other gene, *DMC1*, the most significant SNP was rs763121, with nominal $P = 1.6 \times 10^{-7}$ ($P = 0.0009$ corrected for candidate gene SNP analyses); age at natural menopause was lower by ~0.18 years per copy of the minor allele. *DMC1* encodes a protein that is essential for meiotic homologous recombination and is regulated by *NOBOX*, mutations in which can cause POI[25–27].

Pleiotropy of primary hits

We examined overlap of our significant regions against published GWAS results for other traits (GWAS catalog; see URLs). Twelve menopause loci were within 1 Mb of a previously published genome-wide significant SNP, but most of the colocalized SNPs were in low LD ($0 < r^2 < 0.21$) with our SNP in the region (**Supplementary Table 7**). The exception was at the *GCKR* locus on chromosome 2. Region 2a (rs2303369) was correlated ($r^2 \approx 0.5$) with four different SNPs reported to influence kidney function, type 2 diabetes, continuous glycemic traits, as well as serum albumin, C reactive protein, serum urate, and triglycerides. These results increase the observed clustering of signals in complex trait genetics, whilst also adding to the increasing pleiotropy observed at the *GCKR* locus.

DISCUSSION

In this large two-stage GWAS, we confirmed four established menopause loci and identified and replicated 13 loci newly associated with age at natural menopause. Of these 17 hits, all but two are intronic or exonic to known genes. For associated SNPs in GWAS, on average 40% are intergenic, whereas only 2% of our hits are intergenic. Furthermore, we found twice the nonsynonymous top hits typically observed in GWAS (24% versus 12%; ref. 28). The 17 replicated loci function in diverse pathways including hormonal regulation, immune function and DNA repair. Together, they explained 2.5–4.1% of the population variation in menopausal age in independent replication samples. Biological pathway analysis of the genetic associations with age at natural menopause in this study using distinct algorithms and databases were in close agreement in emphasizing general biological pathways for mitochondrial function, DNA repair, cell cycle and cell death and immune response.

Aging is thought to result from the accumulation of somatic damage[29]. Analysis of gene expression patterns in aging organs, such as heart and brain, identified changes in genes involved in inflammatory response, oxidative stress and genome stability[30], processes also identified in analysis of age-related changes in mouse oocytes, including changes in mitochondrial function[31]. Comparisons of lifespans across species show that longevity and DNA repair function are generally related[32]. This notion is reinforced in the Werner and Bloom syndromes, which involve genome instability due to mutations in 3′→5′ DNA helicases of the RecQ family members, and are characterized by both premature aging and premature menopause[33]. Similarly, an increase in meiotic errors is associated with an age-related decline in oocyte quality, compounding progress toward menopause owing to follicle depletion[34].

In biological pathway analysis, seven candidate genes identified by proximity to the 17 genome-wide significant associations with age at natural menopause are related to DNA damage repair and replication (*EXO1*, *HELQ*, *UIMC1*, *FAM175A*, *FANCI*, *TLK1*, *POLG* and *PRIM1*; **Supplementary Table 4**). The protein encoded by *UIMC1* physically interacts with BRCA1 and estrogen receptor α and is thought to recruit BRCA1 to DNA damage sites and to initiate checkpoint control in the G2/M phase of the cell cycle. PRIM1 (primase) is involved in DNA replication by synthesizing RNA primers for Okazaki fragments during discontinuous replication[35]. A mutation in *POLG* can segregate with POI[36]. *Polg* knock-in mice show lower lifespan, premature aging

and lower fertility compared with wild type[37]. *FANCI*, another gene at the same locus adjacent to *POLG*, is a member of the Fanconi anemia complementation group. Fanconi anemia is a recessive disorder characterized by cytogenetic instability and defective DNA repair. Fanconi anemia patients experience irregular menstruation with menopause occurring around age 30 (ref. 38). The functional polymorphism correlated to our top hit in *EXO1* is associated with longevity in female centenarians[39]. Male and female *Exo1* knockout mice are sterile because the gene is essential for male and female meiosis[40]. In addition to the GWAS regions in or near genes associated with early menopause, we investigated a panel of candidate genes identified before the study, and found a SNP near the meiotic recombination gene *DMC1* to be significantly associated with age at menopause. How the DNA repair pathways contribute to menopause remains unclear. With altered DNA repair mechanisms, damage could accumulate, rendering poor-quality oocytes for selection. In contrast, the number of damaged follicles may increase with aging, leading to a greater rate of follicle loss through atresia. The top hit in this study, a nonsynonymous SNP in *MCM8*, was not included in the IPA results, probably because the exact function of this protein is still unknown. The MCM family, however, is a key component of the prereplication complex, and its main function is to restrict DNA replication to one round per cell cycle[41].

The pathway analyses highlighted additional candidate genes with functions in DNA repair, but with subgenome-wide levels of significance for association with age of natural menopause. These 12 candidates (**Supplementary Table 5**) included the gene encoding Werner helicase (WRN), mutations in which cause Werner syndrome, a classic progeria with advanced aging phenotype and ovarian aging[42]. Estrogen can enhance WRN expression, preventing cell senescence, suggesting that WRN is involved in menopause[43]. The identification of DNA repair as one of the biological pathways involved in menopause may also explain the association between smoking and an earlier age at menopause. Damage caused by smoking activates several different DNA repair mechanisms. Indeed, a polymorphism in *Exo1*, one of our top loci, is associated with colorectal adenomas in smokers only[44]. Additional studies are needed to determine whether smoking status modifies the association between age at natural menopause and polymorphisms in DNA repair genes, as has been observed for various cancers.

Pathway-based analysis indicated that genes related to autoimmune disease also influence age at natural menopause. This link has not been reported before, however, in a proportion (2–10%) of women with POI, ovarian autoimmunity can have a role[45]. POI is frequently associated with additional autoimmune diseases, such as type 1 diabetes mellitus[46]. The top SNP in region 19a is near *IL11*, which binds the interleukin 11 (IL-11) receptor α chain. Female mice with null mutations in *Il11ra* are infertile owing to defective uterine decidualization, the process necessary for successful embryo implantation[47]. *NLRP11* (region 19b) is a member of the NLRP family of genes, which have important roles in the innate immune system and reproductive system. Several NLRP genes show an oocyte-specific expression pattern[46], whereas *NLRP5* has been implicated in POI, and serves as an autoantigen in a mouse model of autoimmune POI[48,49]. Many autoimmune conditions are associated with a particular HLA type, but no such association has been reported for POI[50,51]. One of our top menopause associations (rs1046089) is a missense substitution in *PRRC2A* (HLA-B-associated transcript), which is in the HLA class III complex on chromosome 6 and has been associated with type I diabetes mellitus and rheumatoid arthritis. Multiple phenotypes have been associated with *PRRC2A* SNPs in GWAS, including BMI, neonatal lupus, HIV control and height (**Supplementary Table 7**), but the SNPs have low correlation with our top hit. Expression data for rs1046089 show that the polymorphism was associated with altered expression of HLA-DRB4 in monocytes and HLA-DQA1 in lymphoblastoid cell lines (**Table 2**). Thus, this gene is a candidate for a proinflammatory component to oocyte depletion that affects menopause age. Indeed, the enrichment of genes involved in NF-κB signaling (*TNF*, *TNFRSF17* and *CSNK2B*) in biological pathway analysis suggests that susceptibility to inflammation, which often accompanies immunosenescence in aging, may also affect ovarian aging. The finding that the innate immune response can be upregulated in response to DNA damage[52] suggests that interplay between the two main pathways we identified (DNA repair and inflammation) may contribute to variation in age at natural menopause.

Three of the 17 regions can be linked to hormonal regulation, an additional route to follicle pool exhaustion. The top SNP in region 11 (rs12294104) is in high LD with SNPs in *FSHB* ($r^2 = 0.92$, **Table 2**), which limits the rate of production of FSH, a key pituitary gland–expressed hormone that stimulates maturation of follicles. FSH-deficient female mice are infertile[53]. Transgenic mice that overexpress FSH show premature infertility owing to postimplantation reduction of embryo-fetal survival[54]. FSH concentrations rise in women approaching menopause; this might be related to a decrease in growing follicles[55]. Mutations in *FSHB* cause hypogonadism and primary amenorrhea in women[56] and lead to greater FSH concentrations and infertility in males compared with wild type[57]. The latter observation is due to a promoter polymorphism that may be causal[58] and is in LD ($r^2 = 0.7$) with our most significant SNP. Although *STAR*, which encodes steroidogenic acute regulatory protein (StAR), was not the nearest gene to the top SNP in region 8 (rs2517388), its functional role in cleavage of cholesterol to pregnenolone in response to tropic hormones makes it a probable functional candidate, and our top SNP is in high LD with SNPs in that gene ($r^2 = 0.81$, **Table 2**). Pregnenolone is a precursor for several steroid hormones, such as estrogen and progesterone, and mutations in the *STAR* gene are associated with congenital lipoid adrenal hyperplasia and POI[59]. Furthermore, *STAR* is a target of FOXL2, for which truncating mutations are preferentially associated with POI[60]. Similarly, *BCAR4*, which encodes the breast cancer antiestrogen resistance 4 protein, is the best candidate gene near region 16. *BCAR4* is expressed only in placenta and oocytes and may have a role in hormonal stimulation in the ovary. In breast cancer treatment, tumors highly expressing *BCAR4* are more resistant to tamoxifen treatment[61], reinforcing the role of *BCAR4* in transduction of hormonal signals.

In summary, our findings demonstrate the role of genes that regulate DNA repair and immune function, and genes affecting neuroendocrine pathways of ovarian function in regulating age at menopause, indicating that the process of aging is involved in both somatic and germ line aging.

We expect that several additional common variants with small effects on age at natural menopause are yet to be identified, and that many of them are in genes in pathways identified in this study. Sequencing and exome chip studies to determine whether low-frequency and rare variants of large effect also contribute to age at natural menopause are underway or being planned in many of the cohorts involved in this GWAS. A collaboration of several consortia is examining the contribution of common genetic variants to age at natural menopause in African-American women, and could allow researchers to determine whether the genetic variation that affects age at natural menopause in African-American women is the same or substantially different from that for women of primarily European descent. We are now conducting a study of women with POI to determine whether variants associated with age at natural menopause

within the normal range of age 40–60 also contribute to disease conditions related to the early-menopause phenotype.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are grateful to the study participants and staff from all cohorts involved in this study. Extended acknowledgments per cohort are in **Supplementary Note**.

AUTHOR CONTRIBUTIONS

Individual study design and management were done by D.I.C., C. He, E.M.B., P.K., J.S.B., M. Boban, E.B., H.C., S.J.C., U.d.F., I.J.D., G.V.Z.D., S.E., J.G.E., L. Ferrucci, A.R.F., P.G., C.J.M.W.G., C.G., D.E.G., P.H., S.E.H., A.H., E.I., S.L.R.K., D.A.L., P.K.E.M., M. Marongiu, N.G.M., V.M., N.C.O.-M., G. Paré, A.N.P., N.L.P., P.H.M.P., O.P., B.M.P., K.R., I.R., J.M.S., R.P.S., V.J.M.P., V.G., G.E., T.B.H., L.J.L.,Y.V.S., A.T., Y.T.v.d.S., P.M.V., G. Waeber, H.E.W., J.F.W., B.H.R.W., L.M.L., D.I.B., J.E.B., L.C., F.B.H., D.J.H., G.W.M., B.A.O., P.M.R., D.S., T.D.S., K. Stefansson, E.A.S., M.U., C.M.v.D., H.V., S.B., A. Salumets and A. Metspalu. Data collection was done by B.Z.A., S.B., L.B.L., J.S.B., M. Boban, A.B., H.C., P.dA., U.dF., I.J.D., G.V.Z.D., S.E., J.G.E., L. Ferreli, L. Ferrucci, A.R.F., D.E.G., P.H., S.E.H., A.C.H., A.H., A.C.J.W.J., I.K., S.L., D.A.L., P.K.E.M., N.G.M., I.M.K., A.B.N., N.L.P., P.H.M.P., O.P., B.M.P., K.R., I.R., A. Scuteri, S.N.S., J.M.S., M.G.S., U.S., B.T., L.T., S.U., V.J.M.P., V.G., G.E., Y.T.v.d.S., C.H.v.G., J.M.V., G. Waeber, H.E.W., A.F.W., T.Z., L.Z., M.C.Z., M.Z., L.M.L., A.M.A., J.E.B., E.A.S., A.G.U., J.M.M., P.v.N. and A. Metspalu. Genotyping was done by J.K., L.S., T.E., E.B., F.J.B., S.J.C., P.dA., G.D., P.D., C. Hayward, A.C.H., J. Liu, B.C.J.M.F., S.E.M., V.M., P.N., D.R.N., N.C.O.-M., A.S.P., A.N.P., B.M.P., J.I.R., A. Singleton, K. Stirrups, A.T., K.T., Y.T.v.d.S., M.V., E.W., T.Z., L.M.L., F.B.H., G.W.M., B.A.O. and U.T. Genotype preparation was done by J.R.B.P., M. Barbalic, N.F., E.P., S.-Y.S., W.V.Z., L.B.L., P.dA., G.D., C.J.M.W.G., C.G., C. Hayward, E.I., S.L.R.K., P.K.E.M., I.M.K., C.M., S.E.M., P.N., D.R.N., N.C.O.-M., A.P., G. Pistis, E.R., C.S., J.A.S., H.S., N.S., Y.V.S., A.T., M.V., E.W., T.Z., L.M.L., F.B.H., D.J.H., S.S., K.L.L. and M.H. Phenotype preparation was done by J.R.B.P., L.S., C. He, M. Mangino, M. Barbalic, L.B., E.M.B., F.E., N.F., D.F.G., J.-J.H., P.K., G.Z., W.V.Z., L.B.L., M. Boban, E.J.C.d.G., I.J.D., G.V.Z.D., M.G.S., S.E., J.G.E., C.H.v.G., L. Ferreli, K.F., M.H., C. Hayward, E.I., A.C.J.W.J., S.L.R.K., I.K., J. Lahti, S.L., T.L., D.A.L., L.M.L., P.K.E.M., I.M.K., C.M., P.H.M.P., G. Pistis, O.P., E.R., C.S., A. Scuteri, J.A.S., B.T., S.U., R.M.v.D., V.G., G.E., T.A., Y.T.v.d.S., H.W., G. Willemsen, B.H.R.W., A.F.W., M.C.Z., L.M.L., A.M.A., E.W.D., A. Metspalu, K.L.L. and J.M.M. Analysis plan development was done by L.S., D.I.C., C. He, E.M.B., P.K., T.C., P.G., D.K., D.P.K., N.G.M., D.T., D.J.H., G.W.M., K.L.L., J.M.M. and A. Murray. Analysis plan review was done by L.S., D.I.C., C. He, A.V.S., A.D.C., N.G., D.K., D.P.K., B.M., A.B.N., B.M.P., A.M.A., D.J.H., K.L.L., J.M.M. and A. Murray. Study data analysis was done by L.S., D.I.C., C. He, M. Mangino, M. Barbalic, L.B., E.M.B., F.E., T.E., N.F., D.F.G., J.-J.H., P.K., P.F.M., E.P., S.-Y.S., A.V.S., S.v.W., G.Z., W.V.Z., E.A., C.C., M.C.C., T.C., N.G., T.H., C. Hayward, Z.K., J. Lahti, D.A.L., L.M.L., D.M., N.C.O.M., G. Paré, G. Pistis, A.S.P., V.E., E.R., J.A.S., M.G.S., S.U., Y.T.vdS., M.V., L.M.Y.-A., L.Z., S.S. and U.T. Review and interpretation of analyses were done by J.R.B.P., L.S., D.I.C., C. He, P.S., M. Barbalic, E.M.B., N.F., P.K., P.F.M., A.V.S., B.Z.A., S.J.C., A.D.C., I.J.D., S.E.H., E.I., D.K., S.L.R.K., D.P.K., L.M.L., P.K.E.M., N.G.M., G. Paré, A.S.P., J.A.S., H.S., L.M.Y.-A., J.E.B., E.W.D., F.B.H., D.J.H., G.W.M., P.M.R., C.M.v.D., H.V., K.L.L., J.M.M., P.v.N. and A. Murray. Meta-analyses were done by L.S., P.S. and K.L.L. Pathway and other analyses were done by J.R.B.P., D.I.C., C. He, A.D.J., M. Mangino, G. Paré and J.A.V. Menopause study was designed by J.R.B.P., L.S., D.I.C., C. He, D.E.G., P.H.M.P., Y.T.v.d.S., C.H.v.G., A. Metspalu, K.L.L., J.M.M. and A. Murray. Manuscript was prepared by J.R.B.P., L.S., D.I.C., C. He, T.E., N.F., A.D.C., D.K., D.P.K., E.W.D., A. Metspalu, K.L.L., J.M.M., A. Murray and J.A.V. Manuscript was reviewed by J.R.B.P., L.S., D.I.C., C. He, N.F., P.K., P.F.M., A.V.S., B.Z.A., T.A., L.B.L., J.S.B., M. Boban, F.J.B., H.C., S.J.C., C.C., M.C.C., A.D.C., G.D., U.d.F., I.J.D., G.E., B.C.J.M.F., M.E.G., N.G., D.E.G., P.H., S.E.H., C. Hayward, E.I., D.K., D.P.K., S.L.R.K., J.A.S., I.K., Z.K., T.L., T.E., A. Salumets, A. Metspalu, J.S.E.L., J. Liu, L.M.L., Y.V.L., P.K.E.M., I.M.K., B.M., V.M., P.N., A.B.N., N.C.O.-M., G. Paré, A.N.P., N.L.P., P.H.M.P., O.P., B.M.P., J.I.R., I.R., H.S., J.M.S., R.P.S., A.T., R.M.v.D., Y.T.v.d.S., C.H.v.G., P.M.V., M.V., G. Waeber, J.F.W., B.H.R.W., A.F.W., L.M.Y.-A., T.Z., L.Z., M.C.Z., V.J.M.P., L.M.L., A.M.A., D.I.B., J.E.B., E.W.D., V.G., T.B.H., F.B.H., D.J.H., L.J.L., P.M.R., T.D.S., E.A.S., H.V., K.L.L., J.M.M., A. Murray, P.v.N. and J.A.V. Consortium was overseen by K.L.L., J.M.M., A. Murray and J.A.V.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics/.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Burger, H.G. The menopausal transition. *Baillieres Clin. Obstet. Gynaecol.* **10**, 347–359 (1996).
2. Broekmans, F.J., Soules, M.R. & Fauser, B.C. Ovarian aging: mechanisms and clinical consequences. *Endocr. Rev.* **30**, 465–493 (2009).
3. te Velde, E.R. & Pearson, P.L. The variability of female reproductive ageing. *Hum. Reprod. Update* **8**, 141–154 (2002).
4. te Velde, E.R., Dorland, M. & Broekmans, F.J. Age at menopause as a marker of reproductive ageing. *Maturitas* **30**, 119–125 (1998).
5. Murabito, J.M., Yang, Q., Fox, C., Wilson, P.W. & Cupples, L.A. Heritability of age at natural menopause in the Framingham Heart Study. *J. Clin. Endocrinol. Metab.* **90**, 3427–3430 (2005).
6. Snieder, H., MacGregor, A.J. & Spector, T.D. Genes control the cessation of a woman's reproductive life: a twin study of hysterectomy and age at menopause. *J. Clin. Endocrinol. Metab.* **83**, 1875–1880 (1998).
7. van Asselt, K.M. *et al.* Heritability of menopausal age in mothers and daughters. *Fertil. Steril.* **82**, 1348–1351 (2004).
8. Vink, J. & Boomsma, D.I. Modeling age at menopause. *Fertil. Steril.* **83**, 1068 (2005).
9. He, C. *et al.* A large-scale candidate gene association study of age at menarche and age at natural menopause. *Hum. Genet.* **128**, 515–527 (2010).
10. He, L.N. *et al.* Association study of the oestrogen signalling pathway genes in relation to age at natural menopause. *J. Genet.* **86**, 269–276 (2007).
11. Weel, A.E. *et al.* Estrogen receptor polymorphism predicts the onset of natural and surgical menopause. *J. Clin. Endocrinol. Metab.* **84**, 3146–3150 (1999).
12. Tempfer, C.B. *et al.* Polymorphisms associated with thrombophilia and vascular homeostasis and the timing of menarche and menopause in 728 white women. *Menopause* **12**, 325–330 (2005).
13. van Asselt, K.M. *et al.* Factor V Leiden mutation accelerates the onset of natural menopause. *Menopause* **10**, 477–481 (2003).
14. He, C. *et al.* Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat. Genet.* **41**, 724–728 (2009).
15. Stolk, L. *et al.* Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat. Genet.* **41**, 645–647 (2009).
16. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
17. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
18. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
19. Marini, F. & Wood, R.D. A human DNA helicase homologous to the DNA cross-link sensitivity protein Mus308. *J. Biol. Chem.* **277**, 8716–8723 (2002).
20. Pullirsch, D. *et al.* The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X X inactivation. *Development* **137**, 935–943 (2010).
21. Copeland, W.C. & Longley, M.J. DNA polymerase gamma in mitochondrial DNA replication and repair. *ScientificWorldJournal* **3**, 34–44 (2003).
22. Tschopp, J., Martinon, F. & Burns, K. NALPs: a novel protein family involved in inflammation. *Nat. Rev. Mol. Cell Biol.* **4**, 95–104 (2003).
23. Segrè, A.V. *et al.* Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.* **6,** pii, e1001058 (2010).
24. Raychaudhuri, S. *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.* **5**, e1000534 (2009).
25. Bishop, D.K., Park, D., Xu, L. & Kleckner, N. DMC1: a meiosis-specific yeast homolog of *E. coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69**, 439–456 (1992).
26. Qin, Y. *et al.* NOBOX homeobox mutation causes premature ovarian failure. *Am. J. Hum. Genet.* **81**, 576–581 (2007).
27. Rajkovic, A., Pangas, S.A., Ballow, D., Suzumori, N. & Matzuk, M.M. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science* **305**, 1157–1159 (2004).
28. Manolio, T.A. Genomewide association studies and assessment of the risk of disease. *N. Engl. J. Med.* **363**, 166–176 (2010).
29. Johnson, F.B., Sinclair, D.A. & Guarente, L. Molecular biology of aging. *Cell* **96**, 291–302 (1999).
30. Lee, C.K., Allison, D.B., Brand, J., Weindruch, R. & Prolla, T.A. Transcriptional profiles associated with aging and middle age-onset caloric restriction in mouse hearts. *Proc. Natl. Acad. Sci. USA* **99**, 14988–14993 (2002).
31. Hamatani, T. *et al.* Age-associated alteration of gene expression patterns in mouse oocytes. *Hum. Mol. Genet.* **13**, 2263–2278 (2004).
32. Promislow, D.E. DNA repair and the evolution of longevity: a critical analysis. *J. Theor. Biol.* **170**, 291–300 (1994).
33. Du, X. *et al.* Telomere shortening exposes functions for the mouse Werner and Bloom syndrome genes. *Mol. Cell. Biol.* **24**, 8437–8446 (2004).

34. Dorland, M., van Kooij, R.J. & te Velde, E.R. General ageing and ovarian ageing. *Maturitas* **30**, 113–118 (1998).
35. Shiratori, A. *et al.* Assignment of the 49-kDa (PRIM1) and 58-kDa (PRIM2A and PRIM2B) subunit genes of the human DNA primase to chromosome bands 1q44 and 6p11.1-p12. *Genomics* **28**, 350–353 (1995).
36. Pagnamenta, A.T. *et al.* Dominant inheritance of premature ovarian failure associated with mutant mitochondrial DNA polymerase gamma. *Hum. Reprod.* **21**, 2467–2473 (2006).
37. Trifunovic, A. *et al.* Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).
38. Wong, J.C. & Buchwald, M. Disease model: Fanconi anemia. *Trends Mol. Med.* **8**, 139–142 (2002).
39. Nebel, A. *et al.* A functional EXO1 promoter variant is associated with prolonged life expectancy in centenarians. *Mech. Ageing Dev.* **130**, 691–699 (2009).
40. Wei, K. *et al.* Inactivation of exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes Dev.* **17**, 603–614 (2003).
41. Tye, B.K. Insights into DNA replication from the third domain of life. *Proc. Natl. Acad. Sci. USA* **97**, 2399–2401 (2000).
42. Epstein, C.J., Martin, G.M., Schultz, A.L. & Motulsky, A.G. Werner's syndrome a review of its symptomatology, natural history, pathologic features, genetics and relationship to the natural aging process. *Medicine (Baltimore)* **45**, 177–221 (1966).
43. Lee, S.J., Lee, S.H., Ha, N.C. & Park, B.J. Estrogen prevents senescence through induction of WRN, Werner syndrome protein. *Horm. Res. Paediatr.* **74**, 33–40 (2010).
44. Gao, Y. *et al.* DNA repair gene polymorphisms and tobacco smoking in the risk for colorectal adenomas. *Carcinogenesis* **32**, 882–887 (2011).
45. Welt, C.K. Primary ovarian insufficiency: a more accurate term for premature ovarian failure. *Clin. Endocrinol.* **68**, 499–509 (2008).
46. Hoek, A., Schoemaker, J. & Drexhage, H.A. Premature ovarian failure and ovarian autoimmunity. *Endocr. Rev.* **18**, 107–134 (1997).
47. Robb, L. *et al.* Infertility in female mice lacking the receptor for interleukin 11 is due to a defective uterine response to implantation. *Nat. Med.* **4**, 303–308 (1998).
48. Northup, J., Griffis, K., Hawkins, J., Lockhart, L. & Velagaleti, G. Unusual pseudo dicentric, psu dic (1;19)(q10;q13.42), in a female with premature ovarian failure. *Fertil. Steril.* **87** 697 e5–e8 (2007).
49. Tong, Z.B. & Nelson, L.M. A mouse gene encoding an oocyte antigen associated with autoimmune premature ovarian failure. *Endocrinology* **140**, 3720–3726 (1999).
50. Anasti, J.N. *et al.* Karyotypically normal spontaneous premature ovarian failure: evaluation of association with the class II major histocompatibility complex. *J. Clin. Endocrinol. Metab.* **78**, 722–723 (1994).
51. Jaroudi, K.A., Arora, M., Sheth, K.V., Sieck, U.V. & Willemsen, W.N. Human leukocyte antigen typing and associated abnormalities in premature ovarian failure. *Hum. Reprod.* **9**, 2006–2009 (1994).
52. Gasser, S. & Raulet, D.H. The DNA damage response arouses the immune system. *Cancer Res.* **66**, 3959–3962 (2006).
53. Kumar, T.R., Wang, Y., Lu, N. & Matzuk, M.M. Follicle stimulating hormone is required for ovarian follicle maturation but not male fertility. *Nat. Genet.* **15**, 201–204 (1997).
54. McTavish, K.J. *et al.* Rising follicle-stimulating hormone levels with age accelerate female reproductive failure. *Endocrinology* **148**, 4432–4439 (2007).
55. Lambalk, C.B., De Koning, C.H. & Braat, D.D. The endocrinology of dizygotic twinning in the human. *Mol. Cell. Endocrinol.* **145**, 97–102 (1998).
56. Kottler, M.L. *et al.* A new FSHβ mutation in a 29-year-old woman with primary amenorrhea and isolated FSH deficiency: functional characterization and ovarian response to human recombinant FSH. *Eur. J. Endocrinol.* **162**, 633–641 (2010).
57. Grigorova, M. *et al.* Increased prevalance of the -211 T allele of follicle stimulating hormone (FSH) β subunit promoter polymorphism and lower serum FSH in infertile men. *J. Clin. Endocrinol. Metab.* **95**, 100–108 (2010).
58. Hoogendoorn, B. *et al.* Functional analysis of human promoter polymorphisms. *Hum. Mol. Genet.* **12**, 2249–2254 (2003).
59. Kaku, U. *et al.* Ovarian histological findings in an adult patient with the steroidogenic acute regulatory protein (StAR) deficiency reveal the impairment of steroidogenesis by lipoid deposition. *Endocr. J.* **55**, 1043–1049 (2008).
60. Pisarska, M.D., Bae, J., Klein, C. & Hsueh, A.J. Forkhead l2 is expressed in the ovary and represses the promoter activity of the steroidogenic acute regulatory gene. *Endocrinology* **145**, 3424–3433 (2004).
61. Godinho, M., Meijer, D., Setyono-Han, B., Dorssers, L.C. & Agthoven, T.V. Characterization of BCAR4, a novel oncogene causing endocrine resistance in human breast cancer cells. *J. Cell. Physiol.* **226**, 1741–1749 (2011).

Lisette Stolk[1,2,117], John R B Perry[3,4,117], Daniel I Chasman[5,6], Chunyan He[7,8], Massimo Mangino[9], Patrick Sulem[10], Maja Barbalic[11], Linda Broer[12], Enda M Byrne[13], Florian Ernst[14], Tõnu Esko[15–17], Nora Franceschini[18], Daniel F Gudbjartsson[10], Jouke-Jan Hottenga[19], Peter Kraft[20,21], Patrick F McArdle[22], Eleonora Porcu[23], So-Youn Shin[24], Albert V Smith[25,26], Sophie van Wingerden[12], Guangju Zhai[9,27], Wei V Zhuang[28], Eva Albrecht[29], Behrooz Z Alizadeh[30], Thor Aspelund[25,26], Stefania Bandinelli[31], Lovorka Barac Lauc[32], Jacques S Beckmann[33,34], Mladen Boban[35], Eric Boerwinkle[11], Frank J Broekmans[36], Andrea Burri[9], Harry Campbell[37], Stephen J Chanock[38], Constance Chen[20,39], Marilyn C Cornelis[39], Tanguy Corre[40], Andrea D Coviello[41,42], Pio d'Adamo[43,44], Gail Davies[45], Ulf de Faire[46], Eco J C de Geus[19,47], Ian J Deary[45,48], George V Z Dedoussis[49], Panagiotis Deloukas[24], Shah Ebrahim[50], Gudny Eiriksdottir[25], Valur Emilsson[25], Johan G Eriksson[51–55], Bart C J M Fauser[36], Liana Ferreli[23], Luigi Ferrucci[56], Krista Fischer[15], Aaron R Folsom[57], Melissa E Garcia[58], Paolo Gasparini[43,44], Christian Gieger[29], Nicole Glazer[41], Diederick E Grobbee[59], Per Hall[60], Toomas Haller[15], Susan E Hankinson[20,61], Merli Hass[15], Caroline Hayward[62], Andrew C Heath[63], Albert Hofman[2,12], Erik Ingelsson[60], A Cecile J W Janssens[12], Andrew D Johnson[42], David Karasik[42,64], Sharon L R Kardia[65], Jules Keyzer[66], Douglas P Kiel[42,64], Ivana Kolcic[35], Zoltán Kutalik[33,67], Jari Lahti[68], Sandra Lai[23], Triin Laisk[69], Joop S E Laven[70], Debbie A Lawlor[71], Jianjun Liu[72], Lorna M Lopez[45,48], Yvonne V Louwers[70], Patrik K E Magnusson[60], Mara Marongiu[23], Nicholas G Martin[13], Irena Martinovic Klaric[73], Corrado Masciullo[40], Barbara McKnight[74], Sarah E Medland[13], David Melzer[3], Vincent Mooser[75], Pau Navarro[62], Anne B Newman[76], Dale R Nyholt[13], N Charlotte Onland-Moret[59], Aarno Palotie[24,77,78], Guillaume Paré[5,6,79], Alex N Parker[80,81], Nancy L Pedersen[60], Petra H M Peeters[59,82], Giorgio Pistis[40], Andrew S Plump[83], Ozren Polasek[35], Victor J M Pop[84], Bruce M Psaty[85,86], Katri Räikkönen[68], Emil Rehnberg[60], Jerome I Rotter[87], Igor Rudan[35,37], Cinzia Sala[40], Andres Salumets[15,69,88], Angelo Scuteri[89], Andrew Singleton[90], Jennifer A Smith[65], Harold Snieder[30,91], Nicole Soranzo[9,24], Simon N Stacey[10], John M Starr[48,92], Maria G Stathopoulou[49,93], Kathleen Stirrups[24], Ronald P Stolk[30,91], Unnur Styrkarsdottir[10], Yan V Sun[94], Albert Tenesa[62,95], Barbara Thorand[96], Daniela Toniolo[40,97], Laufey Tryggvadottir[26,98], Kim Tsui[80], Sheila Ulivi[43], Rob M van Dam[39,99,100], Yvonne T van der Schouw[59], Carla H van Gils[59], Peter van Nierop[101], Jacqueline M Vink[19], Peter M Visscher[48,102], Marlies Voorhuis[36,59], Gérard Waeber[103], Henri Wallaschofski[104], H Erich Wichmann[105–107], Elisabeth Widen[77], Colette J M Wijnands-van Gent[108], Gonneke Willemsen[19], James F Wilson[37], Bruce H R Wolffenbuttel[91,109], Alan F Wright[62], Laura M Yerges-Armstrong[22],

Tatijana Zemunik[35], Lina Zgaga[37,110], M Carola Zillikens[1], Marek Zygmunt[111], The LifeLines Cohort Study[91], Alice M Arnold[74], Dorret I Boomsma[19,47], Julie E Buring[5,6,112], Laura Crisponi[23], Ellen W Demerath[57], Vilmundur Gudnason[25,26], Tamara B Harris[58], Frank B Hu[20,39,61], David J Hunter[20,21,39,61], Lenore J Launer[58], Andres Metspalu[15–17,88], Grant W Montgomery[13], Ben A Oostra[113], Paul M Ridker[5,6,112,114], Serena Sanna[23], David Schlessinger[115], Tim D Spector[9], Kari Stefansson[10,26], Elizabeth A Streeten[22], Unnur Thorsteinsdottir[10,26], Manuela Uda[23], André G Uitterlinden[1,2,12], Cornelia M van Duijn[2], Henry Völzke[116], Anna Murray[3,117], Joanne M Murabito[41,42,117], Jenny A Visser[1,117] & Kathryn L Lunetta[28,42,117]

[1]Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. [2]Netherlands Consortium of Healthy Aging, Rotterdam, The Netherlands. [3]Peninsula Medical School, University of Exeter, Exeter, UK. [4]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [5]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [6]Harvard Medical School, Boston, Massachusetts, USA. [7]Department of Public Health, Indiana University School of Medicine, Indianapolis, Indiana, USA. [8]Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, Indiana, USA. [9]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [10]deCODE Genetics, Reykjavik, Iceland. [11]Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA. [12]Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands. [13]Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [14]Interfakultäres Institut für Genomforschung, Universität Greifswald, Greifswald, Germany. [15]Estonian Genome Center, University of Tartu, Tartu, Estonia. [16]Estonian Biocenter, Tartu, Estonia. [17]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [18]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [19]Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands. [20]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [21]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [22]Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA. [23]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Cagliari, Italy. [24]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. [25]Icelandic Heart Association, Kopavogur, Iceland. [26]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [27]Discipline of Genetics, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada. [28]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. [29]Institute of Genetic Epidemiology, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany. [30]Department of Epidemiology, University Medical Center Groningen, University of Groningen, The Netherlands. [31]Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy. [32]Croatian Science Foundation, Zagreb, Croatia. [33]Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland. [34]Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV), University Hospital, Lausanne, Switzerland. [35]Faculty of Medicine, University of Split, Split, Croatia. [36]Department of Reproductive Medicine and Gynaecology, University Medical Center Utrecht, Utrecht, The Netherlands. [37]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK. [38]Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health (NIH), Bethesda, Maryland, USA. [39]Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA. [40]Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy. [41]Sections of General Internal Medicine, Preventive Medicine and Epidemiology, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA. [42]National Heart, Lung and Blood Institute Framingham Heart Study, Framingham, Massachusetts, USA. [43]Institute for Maternal and Child Health, Istituto di Ricovero e Cura a Carattere Scientifico Materno-infantile Burlo Garofolo Trieste, Trieste, Italy. [44]University of Trieste, Trieste, Italy. [45]Department of Psychology, The University of Edinburgh, Edinburgh, UK. [46]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [47]EMGO Institute for Health and Care Research, VU Medical Center, Amsterdam, The Netherlands. [48]Centre for Cognitive Aging and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, UK. [49]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece. [50]Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK. [51]National Institute for Health and Welfare, Helsinki, Finland. [52]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland. [53]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland. [54]Folkhalsan Research Centre, Helsinki, Finland. [55]Vasa Central Hospital, Vasa, Finland. [56]Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging (NIA), Baltimore, Maryland, USA. [57]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA. [58]Laboratory of Epidemiology, Demography and Biometry, NIA, NIH, Bethesda, Maryland, USA. [59]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands. [60]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [61]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [62]Medical Research Council (MRC) Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK. [63]Washington University St. Louis, St. Louis, Missouri, USA. [64]Hebrew SeniorLife Institute for Aging Research, Harvard Medical School, Boston, Massachusetts, USA. [65]Department of Epidemiology, University of Michigan, Ann Arbor, Michigan, USA. [66]Diagnostic GP Laboratory Eindhoven, Eindhoven, The Netherlands. [67]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [68]Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland. [69]Department of Obstetrics and Gynecology, University of Tartu, Tartu, Estonia. [70]Division of Reproductive Medicine, Department of Obstetrics and Gynaecology, Erasmus MC, Rotterdam, The Netherlands. [71]MRC Centre for Causal Analysis in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, UK. [72]Human Genetic, Genome Institute of Singapore, Singapore. [73]Institute for Migration and Ethnic Studies, Zagreb, Croatia. [74]Department of Biostatistics, University of Washington, Seattle, Washington, USA. [75]Genetics Division, GlaxoSmithKline, King of Prussia, Pennsylvania, USA. [76]Departments of Epidemiology and Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania. [77]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland. [78]Department of Medical Genetics, Helsinki University Central Hospital, Helsinki, Finland. [79]Genetic and Molecular Epidemiology Laboratory, McMaster University, Hamilton, Ontario, Canada. [80]Amgen, Cambridge, Massachusetts, USA. [81]Foundation Medicine, Cambridge, Massachusetts, USA. [82]Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London, UK. [83]Cardiovascular Diseases, Merck Research Laboratory, Rahway, New Jersey, USA. [84]Department of Clinical Health Psychology, University of Tilburg, Tilburg, The Netherlands. [85]Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, Washington, USA. [86]Group Health Research Institute, Group Health Cooperative, Seattle, Washington, USA. [87]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [88]Competence Centre on Reproductive Medicine and Biology, Tartu, Estonia. [89]Lab Cardiovascular Sciences, NIA, NIH, Baltimore, Maryland, USA. [90]Laboratory of Neurogenetics, NIA, Bethesda, Maryland, USA. [91]LifeLines Cohort Study and Biobank, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [92]Geriatric Medicine Unit, University of Edinburgh, Edinburgh, UK. [93]Cardiovascular Genetics Research Unit, Université Henri Poincaré, Nancy, France. [94]Department of Epidemiology, Emory University, Atlanta, Georgia, USA. [95]Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Roslin, UK. [96]Institute of Epidemiology II, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany. [97]Institute of Molecular Genetics, Consiglio Nazionale delle Ricerche, Pavia, Italy. [98]Icelandic Cancer Registry, Reykjavik, Iceland. [99]Saw Swee Hock School of Public Health, National University of Singapore, Singapore. [100]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [101]Municipal Health Service Brabant-Zuidoost, Helmond, The Netherlands. [102]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [103]Department of Internal Medicine, CHUV, University Hospital, Lausanne, Switzerland. [104]Institute for Clinical Chemistry and Laboratory Medicine, University of Greifswald, Greifswald, Germany. [105]Institute of Epidemiology I, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany. [106]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [107]Klinikum Grosshadern, Munich, Germany. [108]POZOB Veldhoven, Veldhoven, The Netherlands. [109]Department of Endocrinology, University Medical Center Groningen, University of Groningen, The Netherlands. [110]Andrija Stampar School of Public Health, Medical School, University of Zagreb, Zagreb, Croatia. [111]Klinik für Gynäkologie und Geburtshilfe, Universität Greifswald, Greifswald, Germany. [112]Harvard School of Public Health, Boston, Massachusetts, USA. [113]Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands. [114]Division of Cardiology, Brigham and Women's Hospital, Boston, Massachusetts, USA. [115]NIA, Intramural Research Program, Baltimore, Maryland, USA. [116]Institut für Community Medicine, Universität Greifswald, Greifswald, Germany. [117]These authors contributed equally to this work. Correspondence should be addressed to A.M. (anna.murray@pms.ac.uk) or K.L.L. (klunetta@bu.edu).

ONLINE METHODS

Discovery. Age at natural menopause was defined as age at last menstrual period that occurred naturally with at least 12 consecutive months of amenorrhea. This analysis included women with natural menopause aged 40–60. Women of self-reported non-European ancestry were excluded, as were women with menopause owing to hysterectomy and/or bilateral ovariectomy, chemotherapy or irradiation, if validated by medical records, and women using HRT before menopause. Most cohorts collected age at natural menopause retrospectively; in the Framingham Offspring, the Atherosclerosis Risk in Communities Study (ARIC), Nurses Health Study (NHS) and WGHS studies, some women became menopausal under study observation. Study-specific questions, mean age at menopause and age at interview are in **Supplementary Table 1**. Genotyping and imputation information for discovery cohorts are in **Supplementary Table 2**. Descriptions of each study are in **Supplementary Note**. For all studies participating in the meta-analyses, each participant provided written informed consent. The Institutional Review Board at the parent institution for each respective cohort approved the study protocol.

Replication. A total of 14,435 women from 21 studies meeting the same inclusion and exclusion criteria as the women in discovery analysis were included in replication analysis. The women had mean and s.d. of age at natural menopause similar to the discovery set (**Supplementary Table 1**). Genotyping and imputation methods for the *in silico* replication cohorts are in **Supplementary Table 2**. Genotyping information for the studies that genotyped the SNPs *de novo* is in **Supplementary Table 2**. Descriptions of each study are in **Supplementary Note**.

The 19 independent genome-wide significant SNPs were tested for association with age at natural menopause using linear regression models. Meta-analysis inverse variance weighted meta-analysis of the studies was done with METAL using genomic control[62]. A SNP within a study was omitted if the minor allele frequency was <1% or imputation quality score was <0.2. The discovery meta-analysis included 2,551,160 autosomal SNPs and 38,968 samples.

Expression quantitative trait locus analysis. For each of the genome-wide significant menopause SNPs (**Table 1**), all proxy SNPs with $r^2 > 0.8$ were determined in HapMap CEU release 22. Each SNP and its proxies were searched against a collected database of expression quantitative trait locus (eQTL) results, including the following tissues: fresh lymphocytes[63], fresh leukocytes[64], leukocyte samples in individuals with Celiac disease[65], lymphoblastoid cell lines (LCLs) derived from asthmatic children[66], HapMap LCLs from three populations[67], HapMap CEU LCLs[68], fibroblasts, T cells and LCLs derived from cord blood[69], peripheral blood monocytes[70,71], $CD4^+$ lymphocytes[72], adipose and blood samples[73], brain cortex[70,74], brain regions including prefrontal cortex, visual cortex and cerebellum (three large studies; V.E., unpublished data), cerebellum, frontal cortex, temporal cortex and caudal pons[75], prefrontal cortex[76], liver[77] and osteoblasts[78]. The collected eQTL results met criteria for statistical significance for association with gene transcript levels as described in the original papers. eQTL findings for replicated GWAS SNPs are summarized in **Table 2**.

Conditional analysis. On each chromosome, SNPs of the lowest *P* value that met genome-wide significance were identified. Genome-wide-significant SNPs >250,000 bp and <1 Mb apart that also had pairwise HapMap CEU LD values of $r^2 < 0.5$ were considered potentially independent regions. Potential independent regions that were within 1 Mb of a second region with a more significant *P* value were tested for independence using conditional analysis. In this analysis, the most significant SNP in the most significant region on each chromosome was used as a covariate in a genome-wide analysis. The second region on the chromosome was then retested for independent association.

Pathway analyses. IPA Knowledge Base 8.8 (see URLs) was used to explore the functional relationship between proteins encoded by the 17 replicated menopause loci. The IPA Knowledge Base contains millions of findings curated from the literature. All reference genes ($n = 61$) within 60 kb potentially encoded by the 17 loci (**Table 2**) were entered into the Ingenuity database. Fifty-one genes were eligible for pathway analysis. These eligible 'focus genes' were analyzed for direct interactions only. Networks were generated with a maximum size of 35 genes and molecular relationships between genes or gene products were graphically represented. Proteins are depicted as nodes in various shapes representing functional class of protein. Lines depict the biological relationships between nodes. To determine the probability that the analyzed gene would be found in a network from Ingenuity Pathways Knowledge Base owing to random chance alone, IPA applies a Fisher's exact test. The network score or *P* value represents the significance of the focus gene enrichment. Enrichment of focus genes to diseases and functional categories was also evaluated in the IPA Knowledge Base. The *P* value, determined by a right-tailed Fisher's exact test, considers the number of identified focus genes and the total number of molecules known to be associated with these categories in the IPA knowledge database.

MAGENTA was used to explore pathway-based associations in the full GWAS data set. MAGENTA implements a GSEA-based approach that has been described[23]. Briefly, each gene in the genome is mapped to a single index SNP of the lowest *P* value within a 110 kb–upstream, 40 kb–downstream window. This *P* value, representing a gene score, is then corrected for confounding factors such as gene size, SNP density and LD-related properties in a regression model. Genes within the HLA region were excluded from analysis owing to difficulties in accounting for gene density and LD patterns. Each mapped gene in the genome is then ranked by its adjusted gene score. At a given significance threshold (95th and 75th percentiles of all gene scores), the observed number of gene scores in a given pathway, with a ranked score above the specified threshold percentile, is calculated. This observed statistic is then compared with 1,000,000 randomly permuted pathways of identical size. This generates an empirical GSEA *P* value for each pathway. Significance was determined when an individual pathway reached FDR <0.05 in either analysis (**Supplementary Table 5**). In total, 2,580 pathways from Gene Ontology, PANTHER, KEGG and Ingenuity were tested for enrichment of multiple modest associations with age at natural menopause.

GRAIL is designed to provide evidence for related biological function among a set of candidate genes. The method is based on connections between gene names and informative words extracted from PubMed abstracts by automated language processing techniques. Genes are considered related, and achieve a high similarity score, if they share informative words. For this analysis, the input for GRAIL was a list of candidate SNPs associated with age at natural menopause. From among candidate genes mapping near the candidate SNPs, GRAIL identifies genes with associated informative words that are significantly similar to informative words from other candidate genes. Genes with significant similarity scores are thus consistent with the set of candidate genes as a whole in having greater sharing of informative words than would be expected by chance, suggesting shared biological functions or even biological pathways. GRAIL was first applied to the lead SNPs from each of the replicating genome-wide significant loci using the 2006 edition of the database of genes and informative words. Separately, GRAIL was applied to a list of 66 SNPs, one from each locus that had at least one SNP meeting a FDR threshold of 0.05 from the QVALUE software in R[79]. For meta-analysis of age at natural menopause, the FDR < 0.05 threshold implied $P < 2.8 \times 10^{-5}$.

Candidate gene analysis. We explored the association of natural age of menopause with 125 candidate genes selected because of a reported relationship with ovarian function, including animal models in which gene mutations affect ovarian function ($n = 37$), human studies of menopause or isolated POI ($n = 48$), syndromes including ovarian failure ($n = 4$) or genes expressed in the ovary or female germ cells ($n = 38$; **Supplementary Table 6**). For each gene, the start and end of transcription was defined by the transcripts that span the largest portion of the genome. NCBI36/hg18 positions taken from the UCSC genome browser were used to define gene and SNP locations. Using the correlation measured from a set of ~850 independent Framingham Heart Study participants, we computed the effective number of independent SNPs for each chromosome[80], and used the total (5,774) in a Bonferroni correction for multiple testing.

Pleiotropy of primary hits. Allelic pleiotropy was explored by comparing genome-wide-significant menopause signals to the online catalog of published GWAS (GWAS catalog; see URLs). All reported associations that reached $P < 5 \times 10^{-8}$ and were within 1 Mb of the menopause signal were considered. LD estimates between the SNP pairs were assessed using HapMap (CEU, release 27). Results are in **Supplementary Table 7**.

URLs. GWAS catalog, http://www.genome.gov/gwastudies/; SCAN, http://www.scandb.org/newinterface/about.html/; IPA, http://www.ingenuity.com/.

62. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
63. Göring, H.H. *et al.* Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat. Genet.* **39**, 1208–1216 (2007).
64. Idaghdour, Y. *et al.* Geographical genomics of human leukocyte gene expression variation in southern Morocco. *Nat. Genet.* **42**, 62–67 (2010).
65. Heap, G.A. *et al.* Complex nature of SNP genotype effects on gene expression in primary human leucocytes. *BMC Med. Genomics* **2**, 1 (2009).
66. Dixon, A.L. *et al.* A genome-wide association study of global gene expression. *Nat. Genet.* **39**, 1202–1207 (2007).
67. Stranger, B.E. *et al.* Population genomics of human gene expression. *Nat. Genet.* **39**, 1217–1224 (2007).
68. Kwan, T. *et al.* Genome-wide analysis of transcript isoform variation in humans. *Nat. Genet.* **40**, 225–231 (2008).
69. Dimas, A.S. *et al.* Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* **325**, 1246–1250 (2009).
70. Heinzen, E.L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008).
71. Zeller, T. *et al.* Genetics and beyond–the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* **5**, e10693 (2010).
72. Murphy, A. *et al.* Mapping of numerous disease-associated expression polymorphisms in primary peripheral blood CD4+ lymphocytes. *Hum. Mol. Genet.* **19**, 4745–4757 (2010).
73. Emilsson, V. *et al.* Genetics of gene expression and its effect on disease. *Nature* **452**, 423–428 (2008).
74. Webster, J.A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445–458 (2009).
75. Gibbs, J.R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
76. Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol. Psychiatry* **15**, 779–784 (2010).
77. Schadt, E.E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**, e107 (2008).
78. Grundberg, E. *et al.* Population genomics in a disease targeted primary cell model. *Genome Res.* **19**, 1942–1952 (2009).
79. Dabney, A., Storey, J.D. & Warnes, G.R. qvalue: Q-value estimation for false discovery rate control. R package version 1.22.0. (2010).
80. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**, 221–227 (2005).

doi:10.1038/ng.1051

Social Science & Medicine 74 (2012) 1125–1129

Contents lists available at SciVerse ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Aspirin use and cardiovascular events in social networks

Kate W. Strully [a,b,*], James H. Fowler [c,d], Joanne M. Murabito [e,f], Emelia J. Benjamin [f,g], Daniel Levy [f,h], Nicholas A. Christakis [i,j,k]

[a] Department of Sociology, University at Albany, State University of New York, 1400 Washington Ave, Albany, NY 12222, USA
[b] Department of Epidemiology, School of Public Health, University at Albany, State University of New York, 1400 Washington Ave, Albany, NY 12222, USA
[c] Division of Medical Genetics, University of California, San Diego, La Jolla, CA, USA
[d] Department of Political Science, University of California, San Diego, La Jolla, CA, USA
[e] Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA
[f] Framingham Heart Study, Framingham, MA, USA
[g] Department of Medicine, Boston University School of Medicine, Boston, MA, USA
[h] National Heart, Lung, and Blood Institute, Bethesda, MD, USA
[i] Department of Medicine, Harvard Medical School, Boston, MA, USA
[j] Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
[k] Department of Sociology, Faculty of Arts and Sciences, Harvard University, Cambridge, MA, USA

ARTICLE INFO

Article history:
Available online 4 February 2012

Keywords:
U.S.A.
Aspirin
Prevention
Cardiovascular disease
Social networks

ABSTRACT

We tested whether friends' and family members' cardiovascular health events and also their own aspirin use are associated with the likelihood that an individual takes aspirin regularly. Analyses were based on longitudinal data on 2724 members of the Framingham Heart Study (based in Massachusetts, U.S.A.) who were linked to friends and family members who were also participants in the same study. Men were more likely to take aspirin if a male friend had recently been taking aspirin, and women were more likely to take aspirin if a brother had recently been taking aspirin. Men were also more likely to take aspirin if a brother recently had a cardiovascular event, and women were more likely to take aspirin if a female friend recently experienced a cardiovascular event. Aspirin use is correlated with the health and behavior of friends and family. These findings add to a growing body of evidence which suggests that behavioral changes that promote cardiovascular health may spread through social networks.

Introduction

A large body of evidence shows that aspirin reduces cardiovascular events for patients with and without histories of cardiovascular disease (CVD) (Berger et al., 2006; Campbell, Smyth, Montalescot, & Steinhubl, 2007; Eidelman, Hebert, Weisman, & Hennekens, 2003; Farley, Dalal, Mostashari, & Frieden, 2006). However, several national studies reveal that aspirin is underutilized (Pignone, Anderson, Binns, Tilson, & Weisman, 2007; Stafford, 2000). Increases in aspirin use have been slow given its efficacy, and rates remain low, particularly among outpatients (Stafford & Radley, 2003). Using the 2003 Behavioral Risk Factor Surveillance Survey, Ajani, Ford, Greenland, Giles, and Mokdad (2006) report adjusted aspirin prevalence rates of 69.3% and 32.7% for individuals with and without CVD. They also find less aspirin use among women than men (adjusted prevalence for women was 34% and for men 38.5%).

Given that aspirin is inexpensive, available without a prescription, and its health benefits have been relatively widely publicized (e.g., industry-initiated advertising campaigns), much of the variation in aspirin use is likely to be determined by factors outside of the clinical setting. While having had a conversation with one's doctor about aspirin is an important determinant of aspirin use in national samples (Brown et al., 2002; Pignone et al., 2007), discrepancies between medical records and self-reports suggest than many people take aspirin on their own initiative without their doctor's knowledge (Brown et al., 2002).

It seems reasonable to suppose that having a member of one's social circle (i.e., an "alter") experience a cardiovascular event and/or begin taking aspirin is likely to affect an individual's (i.e., an "ego") aspirin use. Health behaviors and risk factors (e.g., obesity, smoking) may be "socially contagious," and the chances that one changes his/her own behavior are increased if a member of one's social network recently began behaving differently (e.g., with respect to diet, exercise, smoking, or drinking) (Christakis & Fowler,

* Corresponding author. Department of Sociology, College of Arts and Sciences, University at Albany, State University of New York, 1400 Washington Ave, Albany, NY 12222, USA. Tel.: +518 442 4666.
E-mail address: kstrully@albany.edu (K.W. Strully).

0277-9536/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2011.12.033

2007, 2008; Rosenquist, Murabito, Fowler, & Christakis, 2010). Aspirin-use may have a comparable dynamic of social diffusion. Having a friend or family member experience a cardiovascular event and/or begin taking aspirin may increase an ego's odds of aspirin-use by increasing his/her awareness of CVD and aspirin prophylaxis. Alters' events and aspirin-use could also raise ego's subjective assessment of personal cardiovascular risk and improve his/her attitudes toward aspirin use. This work departs from prior work in that it evaluates both the similar behaviors in alters (aspirin use) and also events occurring in alters (e.g., cardiovascular illness) with respect to how they affect ego.

Methods

The Framingham Heart Study Social Networks Study

The Massachusetts, U.S.A. Framingham Heart Study (FHS) began in 1948 with an Original Cohort of 5209 adults. In 1971, the FHS added an Offspring Cohort comprised of 5124 adult children of the Original Cohort and their spouses. In 1994, a minority oversample of 508 people known as the "Omni Cohort" was empaneled. Beginning in 2002, 4095 adults having at least one parent in the Offspring Cohort enrolled in the Third Generation Cohort, along with 103 parents of the Third Generation Cohort participants who were not previously enrolled in the Offspring Cohort. As a means of following participants, the FHS collected regular contact information for participants' close friends and family members. Since the FHS cohorts are family-based and participants are drawn primarily from the Framingham, Massachusetts area, many of the friends and family members that were listed as contacts are also participants in FHS. Connecting participants (i.e., egos) at each wave to their contacts who were also participants (i.e., alters) gives researchers longitudinal data on a network of social connections among participants. Combining these network data with data from periodic physical exams and questionnaires allows us to test how alters' recent events and behaviors impact egos.

The Framingham Heart Study protocol is reviewed by the Boston University Medical Center Institutional Review Board and participants signed written informed consent. Our project was additionally reviewed by the Harvard Medical School Institutional Review Board.

Study sample

Longitudinal data for our analysis come from physical exams and questionnaires performed during three-year periods centered at 1985, 1989, 1992, 1997, and 1999. We limit our analysis to egos who are likely candidates for aspirin prophylaxis—that is, women age 55–70 and men ages 45–79. All egos are from the Framingham Offspring Cohort, while alters may be in any of the Framingham cohorts. In this analysis, we include three types of intimate peer relationships: spousal relationships, sibships, and friendships. We select these relationship types because we expect egos to be more sensitive to events/behaviors among alters they identify with and feel close to. An insufficient number of observations and events made it impossible to examine gender-stratified parent–child relationships.

In the data for our analysis, information on egos' and alters' health events, behaviors, and social ties are time-varying across waves. If a respondent lists a friend as a contact in one wave, but not in the next wave, that tie is dissolved and is not included in subsequent waves. If an alter dies between the last and current wave, we include information about the health events leading to his/her death in the current wave, but that tie will be dissolved and is not included in all subsequent waves. Examining ego-alter ties across all waves, 32% of the observations are spousal ties, 58% are sibling ties, and 11% are friendships.

Study variables

The dependent variable for this analysis is a dichotomous indicator for whether ego is taking aspirin on a daily basis at the time of the current wave. The key predictors of interest are whether alter was taking aspirin on a daily basis at the last wave and whether alter had a cardiovascular event between the last and current waves. Specific cardiovascular events included in the alter event measure are: myocardial infarction, angina pectoris, coronary insufficiency, stroke, intermittent claudication, and death from CVD or stroke.

We include a lagged version of the dependent variable as a predictor to adjust for whether ego was taking aspirin at the last wave. We also adjust for whether ego has ever had his/her own cardiovascular event prior to the current wave. The measure of ego's prior events includes: myocardial infarction, angina pectoris, coronary insufficiency, stroke, and intermittent claudication. Other control variables are ego's age, education, marital status, and survey wave.

Statistical analysis

The longitudinal logistic regression model employed in our analysis can be written as:

$$\ln\left(\frac{Y^e_{it}}{1-Y^e_{it}}\right) = \alpha + \beta_1 X^a_{1ir(t,t-1)} + \beta_2 X^a_{2ir(t-1)} + \beta_3 X^e_{3i(t-1)} + \beta_4 X^e_{4it} + \beta_5 X^e_{5it} + \beta_6 X_{6t} + e_{irt}$$

where the superscript e indicates a variable measuring an ego characteristic and a superscript a indicates a variable measuring an alter characteristic. The subscript i refers to individual ego, i, the subscript r refers to relationship, r, with a given alter, and the subscript t refers to a given wave at time t. Y^e_{it} is the dichotomous dependent variable indicating whether ego used aspirin at the current wave. $X^a_{1ir(t,t-1)}$ is a dichotomous variable indicating whether alter experienced a cardiovascular health event between the last wave $(t-1)$ and the current wave (t). $X^a_{2ir(t-1)}$ is a dichotomous variable indicating whether alter was taking aspirin at the last wave. $X^e_{3i(t-1)}$ is a dichotomous measure of whether ego was taking aspirin at the last wave. X^e_{4it} is a dichotomous measure of whether ego had a cardiovascular health event at any time prior to the current wave. X^e_{5it} reflects the set of controls for ego's characteristics (e.g., age, education, marital status). X_{6t} reflects the set of controls for survey wave. And, finally, e_{irt} is an error term specific to each ego-alter pair at a given wave. To account for clustered error terms resulting from ego-alter pairings and multiple observations of the same egos across waves, we use generalized estimating equations with an independent working correlations structure (Hardin & Hibe, 2002).

After presenting a model for all ties combined, we stratify the analysis according to ego's sex, the sex-composition of the tie (i.e., two males, two females, or a male and female), and the type of relationship (i.e., spouses, siblings, and friendships). Aspirin-use is lower among women than men, and evidence of the preventative benefits of aspirin is somewhat more controversial for women than men (Ajani et al., 2006; Mulrow & Pignone, 2005). We further expect that people may be more likely to identify with and take behavioral cues from same-gender alters, and that alter's influence on ego may differ depending on the type of relationship they share (i.e., spouses, siblings, friends).

Table 1
Characteristics of egos and alters in sample.

	Total sample	Male egos	Female egos
Ego Aspirin Use at Current Wave	22.98%	25.45%	19.35%
Alter Aspirin Use at Last Wave	20.03%	18.24%	22.66%
Ego Aspirin Use at Last Wave	19.64%	20.98%	17.67%
Alter CV Event between Last and Current Wave	5.23%	4.29%	6.62%
Ego CV Event Prior to Current Wave	12.88%	15.04%	9.70%
Ego Female	40.46%		
Ego Age (average)[a]	61.279 (7.838)	59.333 (8.339)	64.144 (5.978)
Ego Education			
High School Degree	53.99%	47.89%	62.97%
Associates Degree	8.93%	9.69%	7.81%
Bachelor Degree	15.04%	17.14%	11.97%
Masters/Doctorate Degree	10.74%	15.16%	4.23%
None of the Above	11.30%	10.13%	13.01%
Ego Married	85.60%	90.12%	78.96%
N (Ego-Alter Pair Observations)	19,849	11,818	8031

[a] Standard deviations for ego age in parentheses.

Since siblings are, roughly speaking, equally likely to be male as female, there are comparable numbers of same-sex and mixed-sex sibling pairs and we present estimates separately for sister pairs, brother pairs, and brother–sister pairs. Friendships, on other hand, are much more likely to same-sex than mixed-sex, and we do not have sufficient observations to present separate mixed-sex friendship estimates. We, therefore, present estimates for all friendships combined and then present separate estimates for the subset of same-sex friendships. We conducted all the analysis using STATA SE 10 software.

Results

As shown in Table 1, about 25% of male and 19% of female egos were using aspirin on a daily basis in a given wave. Looking at male and female egos combined in the first column of Table 1, about 20% of their alters were using aspirin daily at the last wave and about 5.2% had a cardiovascular event between the previous and current waves. Table 2 presents odds ratios and 95% confidence intervals from our first regression model including all relationship types and both male and female egos. Older egos, male egos, and egos who had a prior cardiovascular event were more likely to take aspirin regularly. In line with widely documented trends, ego's odds of aspirin use were also higher at more recent waves (Stafford, 2000). Ego's marital status and education, however, were not significantly associated with his/her aspirin use in this sample.

In Table 2, ego's odds of daily aspirin use were 12% higher if alter used aspirin daily at the last wave. On the other hand, there was no significant association between alter's recent cardiovascular events and ego's aspirin use. However, this reflects the overall association for all ego-alter pairs and we expect associations may vary according to ego's gender, relationship type, and the gender composition of the relationship. Tables 3 and 4 replicate the basic model presented in Table 2 stratified according to each of these factors.

In Table 3, male egos' aspirin-use was not associated with their wives' aspirin-use or cardiovascular events. Their aspirin use also did not appear to be sensitive to their sisters' aspirin-use or events. On the other hand, male egos were about 41% more likely to use aspirin if their brothers had a cardiovascular event since the last wave. Brothers' earlier aspirin use, however, was not associated with ego's aspirin use. Men's aspirin-use was about 48% higher if a male friend was using aspirin at the last wave. Men's aspirin-use may also be higher if a male friend had a recent cardiovascular event, but this odds ratio of 1.71 should be interpreted cautiously since the lower bound for the 95% confidence interval is 1.00.

Whereas male egos' aspirin-use was not associated with their sisters' aspirin-use, Table 4 shows that female egos were about 35% more likely to take aspirin if their brother used aspirin at the last wave. Brothers' recent cardiovascular events, however, were not

Table 2
Associations between egos' aspirin use and alters' CV events/aspirin use, odds ratios and 95% confidence intervals for total sample.

	Adjusted odds ratio	95% confidence interval
Alter CV Event between Last and Current Wave	1.08	(0.92, 1.27)
Alter Aspirin Use at Last Wave	1.12	(1.02, 1.24)
Control Variables		
Ego Aspirin Use at Last Wave	7.47	(6.32, 8.82)
Ego CV Event Prior to Current Wave	3.81	(3.11, 4.66)
Ego Age	1.03	(1.02, 1.04)
Ego Female	0.66	(0.57, 0.77)
Ego Education		
Associates Degree	0.97	(0.77, 1.22)
Bachelor Degree	0.92	(0.76, 1.11)
Masters/Doctorate Degree	1.04	(0.83, 1.31)
None of the Above	0.99	(0.79, 1.24)
Ego Married	1.00	(0.82, 1.21)
Survey Wave		
1987–1991	2.25	(1.70, 2.98)
1991–1995	2.05	(1.57, 2.67)
1995–1998	2.38	(1.81, 3.13)
1998–2001	3.30	(2.52, 4.34)
N (Ego-Alter Pair Observations)	19,849	

Table 3
Associations between egos' aspirin use and alters' CV events/aspirin use for male egos, by relationship type.

	Alter's recent cardiovascular event		Alter's previous aspirin use	
	Adjusted odds ratio	95% confidence interval	Adjusted odds ratio	95% confidence interval
Spouses ($N = 3840$)	1.08	(0.68, 1.70)	1.16	(0.93, 1.46)
Opposite-sex Siblings ($N = 3537$)	1.27	(0.79, 2.06)	0.90	(0.69, 1.16)
Same-sex Siblings ($N = 3202$)	1.41	(1.04, 1.93)	1.11	(0.88, 1.40)
All Friends ($N = 1239$)	1.53	(0.90, 2.06)	1.36	(0.96, 1.91)
Same-sex Friends ($N = 1013$)	1.71	(1.00, 2.94)	1.48	(1.03, 2.13)

*Although not included in the interests of space, models for each relationship type include all control variables listed in Table 2 (except Ego Female).

Table 4
Associations between egos' aspirin use and alters' CV events/aspirin use for female egos, by relationship type.

	Alter's recent cardiovascular event		Alter's previous aspirin use	
	Adjusted odds ratio	95% confidence interval	Adjusted odds ratio	95% confidence interval
Spouses ($N = 2482$)	0.93	(0.63, 1.36)	1.30	(1.00, 1.69)
Opposite-sex Siblings ($N = 2257$)	0.67	(0.44, 1.02)	1.35	(1.03, 1.77)
Same-sex Siblings ($N = 2454$)	1.18	(0.62, 2.23)	1.15	(0.88, 1.50)
All Friends ($N = 838$)	1.73	(0.74, 4.05)	0.66	(0.40, 1.11)
Same-sex Friends ($N = 740$)	2.85	(1.27, 6.37)	0.77	(0.45, 1.31)

*Although not included in the interests of space, models for each relationship type include all control variables listed in Table 2 (except Ego Female).

associated with women's aspirin use. Women's odds of aspirin use did not appear to be sensitive to their sisters' aspirin use or health events. Women were more than twice as likely to use aspirin if a female friend recently had a cardiovascular event. However, female friends' recent aspirin use was not associated with women's aspirin use. Women's odds of aspirin use may also be positively associated with their husband's recent aspirin use, however this odds ratio of 1.30 should be interpreted cautiously since the lower bound for the 95% confidence interval is 1.00.

Discussion

Persons in this study were more likely to take aspirin regularly if a friend or family member took aspirin. A growing body of literature shows that behavioral changes that promote cardiovascular health may spread through social networks (Christakis & Fowler, 2007, 2008; Rosenquist et al., 2010). The above results add to this literature by providing preliminary evidence that another important cardiovascular health behavior—regular aspirin use—may also be shaped by how members of one's network behave.

Another contribution of this analysis is the examination of alters' cardiovascular events. This approach has both conceptual and methodological advantages compared to prior work. Conceptually, it broadens the scope of what sorts of inter-personal effects might be relevant to individual and public health. Methodologically, it is advantageous because, when considering the effect of an alter event on an ego behavior, it may reduce concern about confounding (since it is a bit harder to imagine factors that are associated with alter events and ego behaviors than it is to imagine factors that are associated with the same behavior in both egos and alters).

While alters' cardiovascular events were not associated with ego's behavior when looking at average effects among all pairs in Table 2, we did find significant associations among certain subgroups. Men were more likely to take aspirin if their brothers recently had a cardiovascular event, and women were more likely to take aspirin if a same-sex friend recently suffered an event. There has been very little research into how people may learn about their own cardiovascular risk and possibly take cues from their friends or family members' health problems (Khwaja, Sloan, & Chung, 2006). Models and theories of behavioral change posit that health behavior depends on how people perceive their health risk and where they fall on a continuum of "readiness to change" (Prochaska & Velicer, 1997). Having a friend or family member go through an experience like a myocardial infarction or stroke may be an important catalyst on the path toward healthier behaviors and better cardiovascular health.

Because aspirin-use is lower among women than men, and evidence of the benefits of aspirin is somewhat more controversial for women than men (Ajani et al., 2006; Mulrow & Pignone, 2005), we stratified the above models according to ego's sex and the sex-composition of the relationship. While the above analysis does not allow us to draw clear conclusions about sex differences in the spread of aspirin-use through social networks, there are suggestive findings. It is notable, for instance, that only male alters, not female alters, appeared to shape ego's aspirin use. Men's aspirin use was associated with their male friends' aspirin use. Women's aspirin use was associated with their brother's aspirin use. Female alters' (i.e., wives, sisters, female friends) aspirin use, on the other hand, was not significantly associated with ego's aspirin use in any of our stratified models. This greater sensitivity to male alter's aspirin use may result from several different factors including higher rates of aspirin use among men, different perceptions of cardiovascular risk for men and women (Frijling et al., 2004), and gender inequalities in the inter-personal dynamics of ego-alter relationships.

Our analysis also points to possible sex differences in sensitivity to alters' cardiovascular events. Associations between ego's aspirin use and alters' cardiovascular events occurred only within same-sex relationships (i.e., brother pairs for men and same-sex friendships for women). There was no association between the ego's aspirin use and the alter's cardiovascular events within mixed-sex pairs. Although several factors may contribute to this pattern, people probably identify more strongly with, and take more health cues from, people of the same sex.

When interpreting the results of our study, it is important consider the role that subjects' doctors may have played in their behavior. One possible explanation for associations between ego's and alter's aspirin use is that they both have the same physician who encouraged aspirin use. In our study sample, we were able to identify ego-alter pairs who shared the same doctor and adjust our estimates for this potential confounder. Associations between ego's and alter's aspirin use were quite robust to this adjustment and it does not appear that shared doctors can account for our results (see Table S1 in the online data supplement). Unfortunately, our data do not allow us to know whether egos who were influenced by alters' aspirin use involved their doctors in their aspirin use decisions (e.g., when an ego learns of an alter's aspirin use, does she then turn to her physician for advice about aspirin prophylaxis?).

In an effort to focus our study on egos who are likely candidates for aspirin prophylaxis, we analyzed women ages 55–70 and men ages 45–79. We believe this broad, demographically-based definition of an aspirin use risk group is appropriate given that we are examining how people make choices about aspirin use outside of clinical settings. Further, because clinical recommendations regarding regular aspirin use were being developed over the period in our study (i.e., 1971–1998), it is important that we capture a broad segment of the general population.

It should be kept in mind, however, that influences of alters' behaviors and health may operate differently on egos with specific clinical risk factors (e.g., those based on blood pressure, cholesterol, diabetes, etc.). We found consistent results when we replicated our main analysis for the subset of egos with prior cardiovascular events and/or elevated 10-year coronary heart disease risk scores (based on Wilson et al., 1998). However, when we limited the sample further to include only egos with prior cardiovascular events, we no longer found a significant association between egos' and alters' aspirin-use. Interpreting this non-significance is difficult, though, because relatively few egos had prior cardiovascular events, raising concerns about sparse data and large standard errors (these results are presented in the online data supplement). Further research, preferably with larger samples, is needed to know whether individuals at higher risk for cardiovascular disease are more or less sensitive to the health and behaviors of their friends and family. This study makes no suggestions related to clinical guidelines or about who should take regular aspirin doses. Rather, our results document how social networks are correlated with whether or not an individual adopts a regular aspirin regimen, regardless of specific clinical risk factors.

This work has a few notable limitations. It does not randomize individuals into social networks, thus leaving open the possibility that these results may in part reflect homophily-driven selection bias on the basis of unobserved traits (e.g., avidity for drugs) that influence the use of aspirin over time. For instance, we can imagine that individuals with similar tastes, health knowledge, or physical resilience may be more likely to form relationships. Shared leisure time activities, as well as similar sociodemographic positions (e.g., educational backgrounds or professions), may increase the chances that people with underlying tendencies toward aspirin-use meet and form relationships. While observational data can never overcome all concerns about unmeasured confounding, the Framingham

Heart Study Social Networks Study, which is longitudinal and provides several control variables, gives us more leverage on causality than is usually possible with non-experimental data. Most notably, our data allow us to adjust for egos' prior aspirin use and cardiovascular events. Also, as has been widely noted, the Framingham Heart Study sample is somewhat homogenous and does not have a significant percentage of underrepresented minorities.

Because of differences in how questions about aspirin use were asked in earlier and later waves of the Framingham Heart Study, our dependent variable captures daily aspirin use, but does not detect less frequent regular aspirin doses (e.g., taking aspirin every other day). Furthermore, with these data, we can measure egos' aspirin use only every four years or so. Lower levels or shorter-term changes in egos' aspirin use are, therefore, not detected in our study. This may mean that we fail to capture some variation in egos' aspirin use, which would imply our estimates may fall on the conservative side.

Because our dataset captures a limited number of relationship ties, we were unable to stratify the data in certain ways, and there were several questions about heterogeneity across subgroups that we were not able to test. For instance, after stratifying by sex and relationship type, we did not have a sufficient number of observations to test whether associations differed depending on whether ego had cardiovascular disease and/or had taken aspirin previously. It should be kept in mind that the above results reflect average associations for a general population. It remains possible that associations between egos' aspirin use and alters' aspirin use/events may differ when distinguishing between further subgroups. Finally, the analyses comparing results across different types of relationships and exposures are exploratory. We did not have prior hypotheses about effect sizes for different types of relationships or exposures. These results should not be interpreted as evidence of differences in causal effects across groups/exposures. Since there is very little existing research into peer influences in drug-taking behavior, we believe that this type of exploratory analysis provides a useful first step in understanding how networks may shape pharmacotherapy. In this case of an exploratory analysis, it is not clear whether Bonferroni adjustments for multiple hypothesis testing are appropriate; nevertheless, we note that all Bonferroni-adjusted p-values in our analysis were greater than 0.05.

Across a broad swathe of behaviors, people are influenced by those around them. Pharmacotherapy is a behavior, and so we should not be surprised by the fact that people's drug-taking behavior is related to the behavior of those around them, and to the events occurring in those around them. Similar to the person who might stop smoking when his friend gets lung cancer, a person whose friend, sibling, or spouse has a myocardial infarction may be more inclined to take aspirin because he/she now has a palpable demonstration of the occurrences the aspirin is intended to prevent, rather than an abstract admonition. Likewise, those whose friends are taking aspirin might follow suit for a variety of reasons, including the basic realization that taking aspirin is not hard at all. People are connected, and so their health is connected.

Funding sources

This was work was supported by NIH (P-01 AG031093) and by the Pioneer Portfolio of the Robert Wood Johnson Foundation; NHLBI 's Framingham Heart Study is supported by contract number N01-HC-25195.

Disclosures

Two authors (NAC and JHF) have an equity stake in a company, MedNetworks, that is licensed by Harvard and UCSD to apply some of the ideas embodied in this work related to patient and provider networks.

Acknowledgments

The authors thank James O'Malley for statistical advice and Rebecca Joyce and Laurie Meneades for expert assistance in data development.

Appendix. Supplementary material

Supplementary data related to this article can be found online at doi:10.1016/j.socscimed.2011.12.033.

References

Ajani, U., Ford, E., Greenland, K., Giles, W., & Mokdad, A. (2006). Aspirin use among U.S. adults: behavioral risk factor surveillance system. *American Journal of Preventive Medicine, 30*(1), 74–77.

Berger, J., Roncaglioni, M., Avanzini, F., Pangrazzi, I., Tognoni, G., & Brown, D. (2006). Aspirin for the primary prevention of cardiovascular events in women and men: a sex-specific meta-analysis of randomized controlled trials. *Journal of the American Medical Association, 295*(3), 306–314.

Brown, J., Delea, T., Nichols, G., Edelsberg, J., Elmer, P., & Oster, G. (2002). Use of oral antithrombotic agents among health maintenance organization members with atherosclerotic cardiovascular disease. *Archives of Internal Medicine, 162*, 193–199.

Campbell, C., Smyth, S., Montalescot, G., & Steinhubl, S. (2007). Aspirin dose for the prevention of cardiovascular disease: a systematic review. *Journal of the American Medical Association, 297*(18), 2018–2024.

Christakis, N., & Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*(4), 370–379.

Christakis, N., & Fowler, J. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine, 358*(21), 2249–2258.

Eidelman, R., Hebert, P., Weisman, S., & Hennekens, C. (2003). An update on aspirin in the primary prevention of cardiovascular disease. *Archives of Internal Medicine, 163*(17), 2006–2010.

Farley, T., Dalal, M., Mostashari, F., & Frieden, T. (2006). Deaths preventable in the U.S. by improvements in the use of clinical preventative services. *American Journal of Preventive Medicine, 38*(6), 600–609.

Frijling, B., Lobo, C., Keus, I., Jenks, K., Akkermans, R., Hulscher, M., & et al. (2004). Perceptions of cardiovascular risk among patients with hypertension or diabetes. *Patient Education and Counseling, 52*(1), 47–53.

Hardin, J., & Hibe, H. (2002). *Generalized estimating equations*. Chapman and Hall.

Khwaja, A., Sloan, F., & Chung, S. (2006). Learning about individual risk and the decision to smoke. *International Journal of Industrial Organizations, 24*(4), 683–699.

Mulrow, C., & Pignone, M. (2005). An editorial update: should she take aspirin? *Annals of Internal Medicine, 142*(11), 942–943.

Pignone, M., Anderson, G., Binns, K., Tilson, H., & Weisman, S. (2007). Aspirin use among adults aged 40 and older in the United States: results of a national survey. *American Journal of Preventive Medicine, 32*(5), 403–407.

Prochaska, J., & Velicer, W. (1997). The transtheoretical model of health behavioral change. *American Journal of Health Promotion, 12*(1), 38–48.

Rosenquist, J., Murabito, J., Fowler, J., & Christakis, N. (2010). The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine, 152*(7), 426–433.

Stafford, R. (2000). Aspirin use is low among United States outpatients with coronary artery disease. *Circulation, 101*(10), 1097–1101.

Stafford, R., & Radley, D. (2003). The underutilization of cardiac medications of proven benefit, 1990–2002. *Journal of the American College of Cardiology, 41*(1), 56–61.

Wilson, P., D'Agostino, R., Levy, D., Belanger, A., Silbershatz, H., & Kannell, W. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation, 97*, 1837–1847.

Association of Female Reproductive Factors with Body Composition: The Framingham Heart Study

Subbulaxmi Trikudanathan, Alison Pedley, Joseph M. Massaro, Udo Hoffmann, Ellen W. Seely, Joanne M. Murabito, and Caroline S. Fox

Division of Endocrinology, Metabolism, and Diabetes (S.T., E.W.S., C.S.F.), Department of Medicine, Brigham and Women's Hospital, and Department of Radiology (U.H.), Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02115; National Institutes of Health (A.P.), Bethesda, Maryland 20892; Framingham Heart Study (J.M.Ma., J.M.Mu., C.S.F.), Framingham, Massachusetts 01702; Department of Mathematics and Statistics (J.M.Ma.), Boston University, Boston, Massachusetts 02215; Section of General Internal Medicine (J.M.Mu.), Department of Medicine, Boston University School of Medicine, Boston, Massachusetts 02118; and Center for Population Studies (C.S.F.), National Heart, Lung, and Blood Institute, Bethesda, Maryland 20824

Background: Identifying reproductive risk factors in women offers a life course approach to obesity and cardiovascular disease prevention. The association of female reproductive factors with measures of regional body fat distribution has not been comprehensively studied.

Methods: We examined the association of female reproductive factors (age at menarche, parity, age at natural menopause, menopausal status) in association with body composition data from women who participated in the Offspring and the Third Generation Framingham Heart Study cohorts. Visceral adipose tissue (VAT) and sc adipose tissue (SAT) were measured volumetrically by multidetector computerized tomography. We modeled the relationship between each fat depot and female reproductive factors after adjusting for various factors such as age, smoking status, alcohol intake, physical activity index, hormone replacement therapy, and menopausal status.

Results: Earlier age at menarche was associated with increased body mass index (BMI), waist circumference (WC), VAT, and SAT ($P < 0.0001$). This association of earlier menarche with adiposity measures was attenuated after adjusting for BMI (all $P > 0.70$). We observed no association between parity and all parameters of adiposity measurements (all $P > 0.24$). Similarly, age at natural menopause was not associated with measures of body composition. Despite higher mean BMI among the post- (BMI 27.3 kg/m^2) compared with the premenopausal women (BMI 25.9 kg/m^2) in an age-matched analysis, mean VAT was not different between the two groups ($P = 0.30$).

Conclusions: Earlier menarche is associated with overall obesity but not with VAT or SAT after accounting for measures of generalized adiposity. Parity and menopausal age were not associated with adiposity measures. Although postmenopausal women had increased BMI, VAT, and SAT, the association was predominantly due to age. **(*J Clin Endocrinol Metab* 98: 236–244, 2013)**

Cardiovascular disease (CVD) is the leading cause of death in women in the United States (1). Women, when compared with men, may manifest their clinical disease later in life (2) and tend to have less adverse CVD risk factor profiles (3), rendering standard risk prediction algorithms less reliable in women (4). The life course approach allows us to understand the association of female reproductive factors on CVD risk from childhood to later adult life. Female reproductive factors may provide a life course approach to understand-

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

doi: 10.1210/jc.2012-1785 Received March 27, 2012. Accepted September 28, 2012.
First Published Online October 23, 2012

Abbreviations: BMI, Body mass index; CT, computed tomography; CVD, cardiovascular disease; DXA, dual-energy x-ray absorptiometry scan; FHS, Framingham Heart Study; MDCT, multidetector computed tomography; SAT, sc adipose tissue; VAT, visceral adipose tissue; WC, waist circumference.

ing and predicting CVD risk in women at their preclinical stage (5–7).

Although there is accumulating evidence regarding the importance of understanding CVD outcomes related to female reproductive factors, including age at menarche, parity, and age at menopause, less is known regarding their determinants. Generalized obesity is a key risk factor that is associated with several reproductive factors, including age at menarche (8, 9), parity (10), and age at menopause (11). However, the association of body fat distribution and female reproductive factors has been less well characterized. Regional fat distribution, particularly visceral adipose tissue (VAT), is thought to be a unique pathogenic fat depot. VAT has been shown to be associated with hypertension (12), diabetes, insulin resistance (12), and dyslipidemia (13). Understanding whether body fat distribution is associated with female reproductive factors may provide insight into the potential mechanisms linking them to CVD and its risk factors.

Thus, the goal of this study was to examine whether measures of regional body fat distribution, including VAT and sc adipose tissue (SAT) are associated with age of menarche, parity, age of menopause, and menopausal status. We hypothesized that visceral adiposity is associated with earlier menarche, greater parity, and earlier menopause independent of body mass index.

Materials and Methods

Study sample

Participants for the present study were drawn from a subset of women (n = 1638) who participated in the multidetector computed tomography (MDCT) substudy of the Framingham Heart Study (FHS) from 2002 to 2005. These participants were part of the FHS Offspring and Third Generation cohorts. Women aged 40 yr or older, weighing less than 160 kg [owing to weight restrictions for computed tomography (CT) scan machine], and not pregnant, were included in the MDCT substudy (14). All participants in the FHS Offspring and Third Generation cohort underwent a contemporaneous research clinical visit (Offspring examination 7, 1998–2001; Third Generation examination 1, 2002–2005). During this clinic visit, a physical examination along with laboratory analyses was performed. Covariate data were taken from the contemporaneous examination for both cohorts. In addition, the Offspring cohort underwent examinations approximately every 4 yr from study inception in 1971 and reproductive history was updated at each attended examination.

The study protocol was approved by the institutional boards of the Boston University Medical Center and Massachusetts General Hospital. Written consent was obtained from all participants.

Assessment of female reproductive factors

The physician-administered standardized medical history portion of the FHS research examinations included questions on reproductive history. At Offspring examination 2 (1978–1981) and the Third generation examination 1, female participants were asked age at the start of menses and also asked, "How old were you at the time of your first menstrual period (menses)?" Responses were recorded in years. At the Third generation examination 1, women were asked about the number of live births. At all Offspring examinations and at Third generation examination 1, women were asked whether their periods stopped for 1 yr or more and if yes, the cause and age, when periods stopped was recorded. Only women who reported natural menopause were included. Women who reported cessation of menstrual periods due to hysterectomy and/or removal of both ovaries, chemotherapy, radiation, and other causes were excluded. Additionally, women who reported the use of hormone replacement therapy before the cessation of menses were excluded because it can lead to the misclassification of age at menopause.

Adiposity measurements

MDCT participants underwent an eight-slice CT scan of the abdomen in a supine position (Lightspeed Ultra; General Electric, Milwaukee, WI). Twenty-five consecutive 5-mm-thick slices were obtained covering 125 mm above the level of S1. Volumes of SAT and VAT were evaluated as described previously (14). Interreader reproducibility was high, with interclass correlations of 0.992 for VAT and 0.997 for SAT (15). Body mass index (BMI) and waist circumference (WC) were measured at the index examination. BMI was defined as weight (in kilograms) divided by square of height (in meters). WC was measured by trained technicians using a tape measure at the level of the umbilicus with the participant standing erect and arms hanging to the sides.

Covariate assessment

Covariates were ascertained during the contemporaneous FHS examination. Current smokers (as standardized across the Framingham Heart Study) were defined as women who smoked at least one cigarette per day for the year preceding the examination. Alcohol use (as standardized across Framingham Heart Study) was characterized in women as seven or more drinks per week. Physical activity was measured by a questionnaire that recorded the average number of sleep hours and activity hours (sleep, sedentary, slight, moderate, and heavy activity) (16). Physical activity index, a composite score was calculated by adding the weighted time spent in each activity as described in our prior studies (16). At Third generation examination 1, all women were asked, "Have you taken hormone replacement therapy?" The age at which hormone replacement therapy was started and duration of therapy were also noted. Use of estrogen or hormone replacement therapy was also collected in the Offspring examinations. Available data from 592 postmenopausal Offspring women showed 21% currently using estrogen (122 of 592 women). Of those women on estrogen, 55% were also on progesterone.

Statistical analyses

We performed three separate analyses for each female reproductive factor with additional analyses for menopausal status.

Age at menarche

Of the 1638 women in the FHS MDCT substudy, 1622 had both VAT and SAT data available. Age at menarche was available for 1544 women. Of these, 1456 women (41% from Offspring cohort and 59% from Third Generation cohort) had complete covariate and outcome (BMI and WC) information. The accuracy of recall for age at menarche in adulthood has been well documented in several studies (17, 18).

Age at menarche was categorized per 1 yr starting from 9 yr up to 14 yr. Those with ages at menarche at 15 yr old or older were grouped together. The primary analysis incorporated all of the data in continuous form using individual years of age and collapsing at the extremes of age. Women who attained menarche at less than 9 yr of age were excluded from the study. Separate multivariable linear regression models were created to examine the association of continuous age at menarche (exposure) with each measure of adiposity (including BMI, WC, VAT, and SAT as the outcome variables). Covariates included in the primary analyses were age at CT scan, cohort (Offspring *vs.* Third Generation), smoking, alcohol intake, hormone replacement therapy, physical activity, and menopausal status. In secondary analyses, to adjust for overall adiposity, BMI was added to the model as a covariate. To reinforce the analyses, we performed step-wise selection model (with entry and stay cutoff of $P = 0.15$) to determine which of the body composition variables were most associated with age at menarche. All the covariates in the primary analyses were forced into the model after which BMI, WC, VAT, and SAT were allowed to compete and the covariate with the highest partial R-square entering first. We also performed backward elimination (with a stay cutoff of $P < 0.15$). In the backward elimination, all variables were put in and the variable with the lowest partial correlation was removed and the model refit.

Parity

Detailed data on parity was only available in the Third Generation data set. Overall, 900 women in the MDCT substudy were in the Third Generation cohort. All 900 had VAT and SAT data. Of these, 898 women had parity information and 870 women had complete covariate information, resulting in a final sample size of 868 women.

Parity was categorized as no, one, two, three, and four or more live births. Multivariable linear regression models with parity as the exposure and adiposity measures (BMI, WC, VAT, and SAT) as the outcome were performed, adjusting for age at CT scan, smoking, alcohol intake, hormone replacement therapy, physical activity, and menopausal status. In secondary analyses, we additionally adjusted for BMI.

Age at menopause

Of the 1622 women in the FHS MDCT substudy with VAT and SAT data available, across both the Offspring and Third Generation cohorts, 652 women experienced natural menopause. Women with missing age at natural menopause (n = 7), women reporting estrogen use before cessation of menstrual periods (n = 21), and women with primary ovarian insufficiency (natural menopause age <40 yr, n = 28) were excluded. Of the remaining 596 women with body composition measures, 522 had complete covariate and outcome information. Age at menopause was categorized as follows: 40–45, 46–47, 48–49, 50–51, 52–53, and more than 54 yr. Because we had too many individual ages and thus for logistical reasons, we simplified them to age categories. We collapsed the age categories at the extremes into 40–45 yr and 54 yr old or older. Multivariable linear regression models were performed, with age at menopause as the exposure and adiposity measures (BMI, WC, VAT, and SAT) as the outcome. We adjusted for age at CT scan, cohort (Offspring *vs.* Third Generation), smoking, alcohol intake, hormone replacement therapy, and physical activity. In second-order models, we additionally adjusted for BMI. We report our results as least square means and *P* values comparing differences across the groups as obtained from multivariable linear regression models.

Menopausal status

To examine pre- *vs.* postmenopausal body composition, we used two complementary analyses designed to examine the association of menopausal status above that of age. Some previous studies (19, 20) have reported that aging may contribute to body adiposity irrespective of menopausal status.

First, we performed an analysis of the overall sample of women from the Offspring and Third Generation cohort (n = 1280). Of these, 1271 women had age at CT and outcome data and were used for the unmatched model. Of these women, 681 were premenopausal and 590 women were postmenopausal. We modeled measures of body composition (BMI, WC, VAT, and SAT) as a function of menopausal status (before *vs.* after menopause). We used an unadjusted model examining crude differences in adiposity between pre- and postmenopausal women; and additionally adjusted for age at CT scan.

In a secondary, supportive sensitivity analysis, we performed a matched analysis of pre- and postmenopausal women matched by cohort and age at CT scan. Eighty-five pairs were identified matched on age at natural menopausal (63 in Third Generation cohort and 22 in Offspring). Premenopausal women within the same cohort were randomly matched to another postmenopausal woman with the same age at CT. Given that many of the postmenopausal women were older at the time of the CT than women who were premenopausal, the sample was significantly reduced from 1271 to the 85 pairs. We performed paired *t* tests of the differences in body composition and conditional logistic regression models predicting postmenopausal status from VAT and SAT. This conditional model included adjustments for smoking and alcohol intake.

Results

Association between age at menarche and body fat distribution

Cross-sectional characteristics of the study sample by categories of age at menarche are presented in Table 1. In the multivariable linear regression analysis examining the association between age at menarche and body composition (Table 2), we observed an inverse association between menarcheal age and adiposity as measured by BMI, WC, VAT, and SAT (all $P < 0.0001$). For each 1-yr increase in menarche age, VAT was 61.0 cm^3 lower. However, upon adjustment for BMI in the above model, our results were attenuated (β-coefficient 2.78 cm^3, $P = 0.74$). These re-

TABLE 1. Cross-sectional characteristics by categories of age at menarche

	9 yr	10 yr	11 yr	12 yr	13 yr	14 yr	≥15 yr
n (yr)	20	50	187	379	423	200	195
Age (yr)	49.4 (7.8)	52.3 (8.8)	53.1 (10.0)	50.9 (9.5)	51.6 (9.4)	52.2 (10.2)	51.7 (11.0)
Age at CT (yr)	50.6 (8.9)	54.1 (10.0)	55.1 (11.3)	52.6 (10.7)	53.4 (10.8)	54.3 (11.6)	53.3 (12.4)
Current smoker (%)	35.0 (7)	24.0 (12)	8.0 (15)	10.0 (38)	12.1 (51)	12.0 (24)	14.9 (29)
Alcohol intake (%)[a]	10.0 (2)	12.0 (6)	15.5 (29)	12.7 (48)	14.2 (60)	20.5 (41)	14.4 (28)
Physical activity index	37.5 (6.1)	37.0 (6.9)	35.8 (5.3)	36.7 (5.9)	36.4 (5.7)	37.2 (5.5)	37.7 (6.2)
Hormone replacement therapy (%)	45.0 (9)	24.0 (12)	21.9 (53)	19.0 (72)	16.5 (70)	19.0 (38)	11.3 (22)
BMI (kg/m^2)	31.0 (7.4)	28.7 (5.8)	28.6 (6.3)	27.1 (5.6)	27.3 (6.0)	25.5 (5.0)	25.4 (5.1)
WC (cm)	101.0 (17.3)	98.3 (15.7)	96.4 (16.3)	92.7 (15.6)	94.1 (16.0)	89.6 (13.5)	88.9 (13.5)
VAT (cm^3)	1697.3 (1099)	1636 (846)	1504 (866)	1319 (845)	1369 (831)	1263 (800)	1176 (721)
SAT (cm^3)	4022.9 (1841)	3696 (1472)	3556 (1591)	3105 (1474)	3212 (1575)	2842 (1313)	2717 (1410)
Postmenopausal (%)	60.0 (12)	58.0 (29)	59.4 (111)	47.2 (179)	47.0 (199)	47.5 (95)	46.2 (90)

Data are presented as mean (±SD), for categorical data, percentage (n).

[a] Defined as more than seven drinks per week.

sults were confirmed using stepwise regression and backward elimination models. We also found the overall model R^2 was 4.37% with a partial R^2 from BMI of 2.87%. BMI had the largest partial R^2 of all model terms and accounted for more 65% of the total variation explained by the model.

Parity and body fat distribution

Supplemental Table 1, published on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org, contains the descriptive data for 868 women according to their parity status. Table 3 shows the multivariable linear regression model for body composition across various categories of parity. Associations between adiposity measures (BMI, WC, VAT, and SAT) and parity did not reach statistical significance in our study (all $P > 0.24$). For example, for every incremental increase in parity by 1 live birth, VAT was 23.1 cm^3 higher, although this association was not significant ($P = 0.24$).

Age at natural menopause and body fat distribution

Supplemental Table 2 depicts the demographic characteristics of the cohort by categories of age at menopause. BMI ($P = 0.5$), WC ($P = 0.5$), VAT ($P = 0.7$), and SAT ($P = 0.8$) were not associated with categorized menopausal age (Table 4, Supplemental Fig. 1). The adjusted means for VAT in women with menopause age between 48 and 49 yr was 1469 cm^3. VAT was higher for women who had menopause earlier than 48 yr and later than 50 yr, although these trends were not significant ($P = 0.7$). We also performed quadratic model predicting adiposity measures from age at menopause and did not find any relationship with the squared terms.

Menopausal status and body fat distribution

In unadjusted models, postmenopausal women had higher levels of BMI, WC, VAT, and SAT (all $P < 0.0001$, Table 5), which were attenuated upon age adjustment ($P > 0.16$). One notable exception was SAT, which continued to be higher in post- *vs.* premenopausal women (3289.8 *vs.* 2876.2 cm^3) after age adjustment ($P = 0.007$).

To disentangle the associations of chronological age and compare the association of menopausal status alone, additional testing was carried out. We performed an age-matched analysis of pre- and postmenopausal pairs; sample characteristics are displayed in Table 6. The mean difference between the age at CT scan and age at menopause was 2.8 yr (range 2–11 yr). As expected, VAT and SAT

TABLE 2. Multivariable linear regression model[a] associations with body composition from age at menarche[b]

	≤10 yr	11 yr	12 yr	13 yr	14 yr	≥15 yr	Beta estimate for trend (SE)	*P* value for trend[c]
BMI (kg/m^2)	28.1 (0.8)	27.8 (0.5)	26.5 (0.4)	26.8 (0.4)	25.0 (0.5)	25.0 (0.5)	−0.6 (0.1)	<0.0001
WC (cm)	97.1 (2.2)	94.5 (1.3)	92.1 (1.0)	93.1 (1.0)	88.7 (1.2)	88.3 (1.3)	−1.3 (0.3)	<0.0001
VAT (cm^3)	1632 (111)	1478 (66)	1377 (53)	1399 (51)	1284 (63)	1225 (65)	−55.7 (13)	<0.0001
SAT (cm^3)	3572 (215)	3374 (128)	2999 (102)	3097 (98)	2760 (121)	2663 (125)	−136.0 (26)	<0.0001

[a] Model adjusted for age at CT scan, cohort, smoking, alcohol intake, physical activity index, hormone replacement therapy, and menopausal status.

[b] Data represented as least square means with SE.

[c] *P* value for trend after additional adjustment for BMI was attenuated for all the above parameters (all $P > 0.70$).

TABLE 3. Multivariable linear regression model[a] associations and least squares means with body composition from parity[b]

	No births (n = 200)	One birth (n = 109)	Two births (n = 337)	Three births (n = 169)	Four or more births (n = 53)	*P* value for trend[c]
BMI (kg/m^2)	26.1 (0.4)	27.7 (0.5)	26.3 (0.32)	26.4 (0.45)	26.2 (0.8)	0.76
WC (cm)	88.7 (1.0)	93.7 (1.4)	90.6 (0.8)	90.4 (1.1)	90.1 (2.0)	0.55
VAT (cm^3)	1038 (50)	1293 (67)	1114 (38)	1188 (54)	1137 (97)	0.24
SAT (cm^3)	2850 (109)	3307 (147)	2982 (83)	3022 (118)	2983 (211)	0.65

[a] Model adjusted for age at CT scan, cohort, smoking, alcohol intake, physical activity index, hormone replacement therapy, and menopausal status.

[b] Data presented as least square means with SE.

[c] *P* value for trend after BMI adjustment remained nonsignificant.

were slightly higher in post- (VAT: 1251.2 cm^3; SAT: 3213.3 cm^3) *vs.* premenopausal women (VAT: 1123.4 cm^3; SAT: 2872.1 cm^3). However, this difference was not statistically significant ($P = 0.3$ for VAT, $P = 0.19$ for SAT).

Discussion

Principal findings

We observed that earlier age at menarche was associated with increased BMI, WC, VAT, and SAT, although this association was attenuated after adjusting for BMI. Our study did not identify an association between parity and BMI, WC, VAT, and SAT. Similarly, age at natural menopausal was not associated with measures of body composition. We noted that postmenopausal women had increased BMI, WC, VAT, and SAT when compared with premenopausal women; however, this was mostly explained by increased ages among postmenopausal women.

Menarche and body composition

In our study earlier age at menarche was associated with BMI, WC, VAT, and SAT. However, VAT was not associated with earlier menarche after adjusting for BMI. These findings suggest that earlier menarcheal age is associated with generalized but not regional body fat depots.

These findings are consistent with other longitudinal and cross-sectional studies that have shown that early menarche is associated with increased BMI and WC (21, 22). For example, in the European Prospective Investigation into Cancer and Nutrition study (22), per older year of age at menarche, generalized and central adiposity (as measured by waist circumference) was lower. Similarly, the BioCycle study used dual-energy x-ray absorptiometry scan (DXA) to measure total body fat in 253 healthy premenopausal women and found that earlier age of menarche (≤12 yr) was associated with increased BMI and percent total body fat (23). The findings from our study advance the current literature in important ways. Our study used a multidetector CT scan to examine VAT in a large number of women in a contemporary setting. We observed from our data that early age at menarche was associated with generalized adiposity rather than VAT.

Disentangling the associations of childhood adiposity on earlier menarcheal age and subsequent risk for adulthood obesity is challenging. We acknowledge the complex interaction between body fat and puberty (24, 25). Epidemiological studies, including the Bogalusa Heart Study (8), have noted that childhood obesity is associated with both earlier menarcheal age and adulthood obesity. Because childhood obesity influences the timing of menarche, it remains unclear whether childhood obesity simply tracks across to adult obesity (26). Furthermore, many genetic variants associated with the timing of menarche are in or near genes associated with childhood and adulthood BMI, energy homeostasis, and body weight regula-

TABLE 4. Multivariable linear regression model[a] in association with body composition from age at menopause[b]

	40–45 yr (n = 69)	46–47 yr (n = 68)	48–49 yr (n = 103)	50–51 yr (n = 121)	52–53 yr (n = 92)	54+ yr (n = 99)	Global *P* value[c]
BMI (kg/m^2)	27.7 (0.7)	27.7 (0.7)	26.8 (0.6)	27.3 (0.5)	28.2 (0.6)	28.1 (0.6)	0.49
WC (cm)	93.6 (1.9)	94.9 (1.8)	94.8 (1.5)	95.6 (1.4)	95.3 (1.6)	98.2 (1.5)	0.50
VAT (cm^3)	1584 (103)	1492 (102)	1469 (83)	1572 (77)	1622 (88)	1635 (85)	0.71
SAT (cm^3)	3199 (178)	3261 (176)	3173 (143)	3286 (132)	3331 (151)	3459 (147)	0.81

[a] Model adjusted for age at CT scan, cohort, smoking, alcohol intake, physical activity, and hormone replacement therapy.

[b] Data presented as least square means with SE.

[c] *P* value for trend after BMI adjustment remained nonsignificant.

TABLE 5. Comparison of body composition between premenopausal and postmenopausal women using the overall sample (n = 1271)

	Premenopausal women (n = 681)	Postmenopausal women (n = 590)	Unadjusted beta estimate (SE) for menopausal status	Unadjusted *P* value	Beta estimate (SE) for menopausal status after adjusting for age at CT scan	*P* value after adjusting for age at CT
BMI (kg/m^2)	26.1	27.5	1.4 (0.32)	<0.0001	0.8 (0.5)	0.14
WC (cm)	89.2	95.4	6.2 (0.84)	<0.0001	0.6 (1.4)	0.69
VAT (cm^3)	1052	1567	515 (43)	<0.0001	119 (71)	0.09
SAT (cm^3)	2876	3267	390 (85)	<0.0001	434 (142)	0.0023[a]

[a] *P* value significant.

tion (27). For example, *LIN28B* is associated with earlier age at menarche and prolonged increase in BMI in women across adolescence and midadulthood (28). Thus, an important limitation of our study is the lack of data on childhood adiposity, which may confound the association of age at menarche with general adiposity observed in our study.

Parity and body composition

In contrast to much of the prior literature (10, 29, 30), we observed no association between parity and measures of adiposity. However, our findings are supported by results from the Stockholm Pregnancy and Women's Nutrition study. In the Stockholm Pregnancy and Women's Nutrition study, 563 Swedish women were followed up at 1 and 15 yr after the first delivery with questionnaires and anthropometric measurements. At the end of 15 yr, approximately 39% of the women became overweight (BMI > 25 kg/m^2). No difference were noted in terms of parity (2.4 ± 0.8 *vs.* 2.4 ± 0.9 children) between the group that became overweight *vs.* the group that remained normal weight, respectively (31). However, weight gain during pregnancy and high weight retainers at the end of the first year had a greater BMI at 15 yr follow-up. The findings observed by the Coronary Artery Risk Development in Young Adults Study showed that considerable weight gain (3–6 kg) and increased WC (3–6 cm) occurred in women who were overweight before childbirth. In this study, excess weight gain was not associated with higher-order births among parous women. However, increased WC was noted in parous women when compared with nonparous women (5–6 cm, $P < 0.001$) (32).

TABLE 6. Cross-sectional characteristics by menopausal status using an age-matched analysis of 85 pre- and postmenopausal pairs

	Premenopausal	Postmenopausal
Age (yr)	49.3 (3.0)	49.4 (3.0)
Age at CT (yr)	50.5 (2.6)	50.5 (2.6)
Current smoker (%)	10.6 (9)	22.4 (19)
Alcohol intake (%)[a]	15.3 (13)	12.9 (11)
Physical activity Index	36.4 (5.9)	37 (6)
BMI (kg/m^2)	25.9 (5.0)	27.3 (6.5)
WC (cm)	89.2 (13.2)	91.6 (17.1)
VAT (cm^3)	1123 (706)	1251 (815)[b]
SAT (cm^3)	2872 (1339)	3213 (1761)[c]

Data are presented as mean (±SD), for categorical data, percentage (n).

[a] Defined as more than seven drinks per week.

[b] $P = 0.30$ on paired *t* test.

[c] $P = 0.19$ on paired *t* test.

Two studies have specifically focused on parity and visceral adiposity. In a cross-sectional study, 170 nonsmoking Caucasian women underwent DXA and CT scan to assess total body fat and intraabdominal adiposity (33). Parity was associated with VAT ($P = 0.02$) but not with percent body fat ($P = 0.68$), waist circumference ($P = 0.16$), or SAT ($P = 0.98$). Potential reasons for the disparate results from our study could be their relatively small sample size and inclusion of only nonsmokers; several cross-sectional studies have noted that smokers have lower body weight than nonsmokers (34). Our study sample included smokers with a higher prevalence of smoking among nonparous women and women with lower-order births. Current smoking was adjusted for in this analysis.

In the second study (4), 122 premenopausal women had total body fat measured by DXA and VAT measured by CT scan over a 5-yr period. Of these 122 women, only 14 had at least one interim birth. There was no change in total body fat over the 5-yr period for one interim birth *vs.* no interim births (2.7 *vs.* 3.7 kg, $P = 0.55$). However, a larger gain in VAT was observed among women who had at least one interim birth (27.1 *vs.* 9.2 cm^2, $P < 0.01$) (35). Differences with our study include the longitudinal design and the short interval between childbirth and assessment of VAT. It may be plausible that parity does not affect long-term gain in VAT as observed in our study.

Age at menopause and body composition

We did not identify robust associations between menopausal age and BMI, WC, VAT, and SAT. Although there is a large body of literature on menopausal status and body weight, very few studies (11, 36) have examined the as-

sociations of age at menopause and body weight. In a Japanese study (n = 1022), pre- and postmenopausal BMI were studied in three groups of women, depending on the age at the time of menopause (<45 yr, 45–50 yr, > 50 yr). No differences in the BMI were observed among the groups both during early postmenopause and late postmenopause (36). There is a paucity of studies exploring the association of age at the time of menopause with body fat depots.

Menopausal status and body composition

Further analysis to dissect the associations between age and menopausal status in the present study revealed that postmenopausal women had increased BMI, WC, VAT, and SAT compared with premenopausal women, but this association was largely attributable to the age. These findings underscore the important role of aging and the selected indices of body composition.

Several studies have explored the association of chronological aging *vs.* menopause-related changes in the body composition of women (19, 37–39). In the Study of Women's Health Across the Nation study, a cross-sectional study of a large multiethnic cohort of middle-aged women, after adjusting for chronological age, no differences were seen in the BMI between premenopausal women and women who reported natural menopause (40). Similarly, Wing *et al.* (41) prospectively studied 484 middle-aged women over 3 yr and noted a 2.25-kg increase in weight during follow-up; however, this weight was similar among control women who remained premenopausal during the follow-up.

Some studies have evaluated longitudinal changes in body composition among perimenopausal women using imaging techniques. An observational longitudinal study followed up 103 healthy premenopausal women over 4 yr. All women gained SAT over time; however, only women who transitioned through menopause demonstrated a significant increase in VAT (42). In another study, intraabdominal fat by CT scan significantly increased by 21% across menopausal transition (43). However, in this study, no age-matched controls were used. Thus, it remains unclear whether these findings were due to aging (43). In the present work, we have attempted to disentangle the associations of aging compared with menopause transition. In doing so, we show that post- compared with premenopausal women have higher levels of SAT, but levels of VAT are similar. Our results suggest that much of the differences in body composition between pre- and postmenopausal women may be due to aging.

Implications

The overarching aim of our study was to use a life course approach to explore the association of female reproductive factors and body composition. The findings from our study suggest that early menarche in women may be an opportune moment to advocate lifestyle measures to prevent adult obesity. Our study notes that parity and age at menopause *per se* may not be suitable determinants for predicting obesity. However, further studies are necessary to test whether female reproductive factors can be used to target lifestyle interventions in high risk women to prevent the metabolic complications of obesity and CVD.

Strengths and limitations

The strengths of our study include the large population-based cohort without ascertainment for obesity-related conditions. All clinical risk factors were assessed by a study physician. Clinical characteristics were carefully measured and standardized. A highly reproducible MDCT scan was used for accurately measuring visceral adiposity. Limitations of our study include the observational and cross-sectional nature of the data, obscuring our ability to comment on temporality and causality. The study sample was comprised of predominantly Caucasians; thus, the findings may not be generalizable to other ethnic groups. Data on birth weight, prior growth parameters, age at adrenarche, and thelarche were not recorded in this study. Other reproductive factors such as gestational weight gain, polycystic ovarian syndrome, gestational diabetes, and preeclampsia were not examined in this study.

Conclusions

Earlier age at menarche is associated with overall adiposity but is not specific to VAT or SAT. Parity and age at menopause were not associated with parameters of central or generalized adiposity. However, postmenopausal women had higher levels of adiposity, which appears to be predominantly due to age.

Acknowledgments

Address all correspondence and requests for reprints to: Caroline S. Fox, M.D., M.P.H., 73 Mt. Wayte Avenue, Suite no. 2, Framingham, Massachusetts 01702. E-mail: foxca@nhlbi.nih.gov.

This work was partially supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (Contract N01-HC-25195) and Grant T32 HL 007609-26. This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National Institutes of Health and Boston University School of Medicine.

Disclosure Summary: A.P. is an employee of Merck and Co., Inc. The other authors have nothing to disclose.

References

1. **Heron M** 2011 Deaths: leading causes for 2007. Natl Vital Stat Rep 59:1–95
2. **Mosca L, Manson JE, Sutherland SE, Langer RD, Manolio T, Barrett-Connor E** 1997 Cardiovascular disease in women: a statement for healthcare professionals from the American Heart Association Writing Group. Circulation 96:2468–2482
3. **Kuller LH** 2010 Cardiovascular disease is preventable among women. Expert Rev Cardiovasc Ther 8:175–187
4. **Ridker PM, Buring JE, Rifai N, Cook NR** 2007 Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA 297: 611–619
5. **Carr DB, Utzschneider KM, Hull RL, Tong J, Wallace TM, Kodama K, Shofer JB, Heckbert SR, Boyko EJ, Fujimoto WY, Kahn SE** 2006 Gestational diabetes mellitus increases the risk of cardiovascular disease in women with a family history of type 2 diabetes. Diabetes Care 29:2078–2083
6. **Bentley-Lewis R** 2009 Late cardiovascular consequences of gestational diabetes mellitus. Semin Reprod Med 27:322–329
7. **Bellamy L, Casas JP, Hingorani AD, Williams DJ** 2007 Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. BMJ 335:974
8. **Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS** 2003 The relation of menarcheal age to obesity in childhood and adulthood: the Bogalusa Heart Study. BMC Pediatr 3:3
9. **Laitinen J, Power C, Järvelin MR** 2001 Family social class, maternal body mass index, childhood body mass index, and age at menarche as predictors of adult obesity. Am J Clin Nutr 74:287–294
10. **Björkelund C, Lissner L, Andersson S, Lapidus L, Bengtsson C** 1996 Reproductive history in relation to relative weight and fat distribution. Int J Obes Relat Metab Disord 20:213–219
11. **Feng Y, Hong X, Wilker E, Li Z, Zhang W, Jin D, Liu X, Zang T, Xu X, Xu X** 2008 Effects of age at menarche, reproductive years, and menopause on metabolic risk factors for cardiovascular diseases. Atherosclerosis 196:590–597
12. **Vega GL, Adams-Huet B, Peshock R, Willett D, Shah B, Grundy SM** 2006 Influence of body fat content and distribution on variation in metabolic risk. J Clin Endocrinol Metab 91:4459–4466
13. **Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu CY, Vasan RS, Murabito JM, Meigs JB, Cupples LA, D'Agostino Sr RB, O'Donnell CJ** 2007 Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. Circulation 116:39–48
14. **Pou KM, Massaro JM, Hoffmann U, Vasan RS, Maurovich-Horvat P, Larson MG, Keaney Jr JF, Meigs JB, Lipinska I, Kathiresan S, Murabito JM, O'Donnell CJ, Benjamin EJ, Fox CS** 2007 Visceral and subcutaneous adipose tissue volumes are cross-sectionally related to markers of inflammation and oxidative stress: the Framingham Heart Study. Circulation 116:1234–1241
15. **Maurovich-Horvat P, Massaro J, Fox CS, Moselewski F, O'Donnell CJ, Hoffmann U** 2007 Comparison of anthropometric, area- and volume-based assessment of abdominal subcutaneous and visceral adipose tissue volumes using multi-detector computed tomography. Int J Obes (Lond) 31:500–506
16. **Kannel WB, Sorlie P** 1979 Some health benefits of physical activity. The Framingham Study. Arch Intern Med 139:857–861
17. **Casey VA, Dwyer JT, Coleman KA, Krall EA, Gardner J, Valadian I** 1991 Accuracy of recall by middle-aged participants in a longitudinal study of their body size and indices of maturation earlier in life. Ann Hum Biol 18:155–166
18. **Damon A, Bajema CJ** 1974 Age at menarche: accuracy of recall after thirty-nine years. Hum Biol 46:381–384
19. **Wang Q, Hassager C, Ravn P, Wang S, Christiansen C** 1994 Total and regional body-composition changes in early postmenopausal women: age-related or menopause-related? Am J Clin Nutr 60:843–848
20. **Demerath EW, Rogers NL, Reed D, Lee M, Choh AC, Siervogel RM, Chumlea WC, Towne B, Czerwinski SA** 2011 Significant associations of age, menopausal status and lifestyle factors with visceral adiposity in African-American and European-American women. Ann Hum Biol 38:247–256
21. **He C, Zhang C, Hunter DJ, Hankinson SE, Buck Louis GM, Hediger ML, Hu FB** 2010 Age at menarche and risk of type 2 diabetes: results from 2 large prospective cohort studies. Am J Epidemiol 171:334–344
22. **Lakshman R, Forouhi NG, Sharp SJ, Luben R, Bingham SA, Khaw KT, Wareham NJ, Ong KK** 2009 Early age at menarche associated with cardiovascular disease and mortality. J Clin Endocrinol Metab 94:4953–4960
23. **Chen L, Zhang C, Yeung E, Ye A, Mumford SL, Wactawski-Wende J, Schisterman EF** 2011 Age at menarche and metabolic markers for type 2 diabetes in premenopausal women: the BioCycle Study. J Clin Endocrinol Metab 96:E1007–E1012
24. **Baker ER** 1985 Body weight and the initiation of puberty. Clin Obstet Gynecol 28:573–579
25. **Siervogel RM, Demerath EW, Schubert C, Remsberg KE, Chumlea WC, Sun S, Czerwinski SA, Towne B** 2003 Puberty and body composition. Horm Res 60:36–45
26. **Anderson SE, Dallal GE, Must A** 2003 Relative weight and race influence average age at menarche: results from two nationally representative surveys of US girls studied 25 years apart. Pediatrics 111:844–850
27. **Elks CE, Perry JR, Sulem P, Chasman DI, Franceschini N, He C, Lunetta KL, Visser JA, Byrne EM, Cousminer DL, Gudbjartsson DF, Esko T, Feenstra B, Hottenga JJ, Koller DL, Kutalik Z, Lin P, Mangino M, Marongiu M, McArdle PF, Smith AV, Stolk L, van Wingerden SH, Zhao JH, Albrecht E, *et al.* 2010 Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nat Genet 42:1077–1085**
28. **Ong KK, Elks CE, Wills AK, Wong A, Wareham NJ, Loos RJ, Kuh D, Hardy R** 2011 Associations between the pubertal timing-related variant in LIN28B and BMI vary across the life course. J Clin Endocrinol Metab 96:E125–E129
29. **Davis SR** 2000 Identifying and promoting the specific nutrition and physical activity needs of women aged 40 and over. Summary. Med J Aust 173(Suppl):S109
30. **Troisi RJ, Wolf AM, Mason JE, Klingler KM, Colditz GA** 1995 Relation of body fat distribution to reproductive factors in pre- and postmenopausal women. Obes Res 3:143–151
31. **Linné Y, Dye L, Barkeling B, Rössner S** 2003 Weight development over time in parous women—the SPAWN study—15 years follow-up. Int J Obes Relat Metab Disord 27:1516–1522
32. **Gunderson EP, Murtaugh MA, Lewis CE, Quesenberry CP, West DS, Sidney S** 2004 Excess gains in weight and waist circumference associated with childbearing: the Coronary Artery Risk Development in Young Adults Study (CARDIA). Int J Obes Relat Metab Disord 28:525–535
33. **Blaudeau TE, Hunter GR, Sirikul B** 2006 Intra-abdominal adipose tissue deposition and parity. Int J Obes (Lond) 30:1119–1124
34. **Chiolero A, Faeh D, Paccaud F, Cornuz J** 2008 Consequences of smoking for body weight, body fat distribution, and insulin resistance. Am J Clin Nutr 87:801–809
35. **Gunderson EP, Sternfeld B, Wellons MF, Whitmer RA, Chiang V, Quesenberry Jr CP, Lewis CE, Sidney S** 2008 Child-bearing may increase visceral adipose tissue independent of overall increase in body fat. Obesity (Silver Spring, Md) 16:1078–1084
36. **Akahoshi M, Soda M, Nakashima E, Tsuruta M, Ichimaru S, Seto S, Yano K** 2001 Effects of age at menopause on serum cholesterol, body mass index, and blood pressure. Atherosclerosis 156:157–163
37. **Avis NE, Crawford SL** 2001 Menopause and weight. Menopause 8:230–232
38. **Douchi T, Yamamoto S, Yoshimitsu N, Andoh T, Matsuo T, Nagata Y** 2002 Relative contribution of aging and menopause to

changes in lean and fat mass in segmental regions. Maturitas 42: 301–306
39. **Trémollieres FA, Pouilles JM, Ribot CA** 1996 Relative influence of age and menopause on total and regional body composition changes in postmenopausal women. Am J Obstet Gynecol 175:1594–1600
40. **Sternfeld B, Wang H, Quesenberry Jr CP, Abrams B, Everson-Rose SA, Greendale GA, Matthews KA, Torrens JI, Sowers M** 2004 Physical activity and changes in weight and waist circumference in midlife women: findings from the Study of Women's Health Across the Nation. Am J Epidemiol 160:912–922
41. **Wing RR, Matthews KA, Kuller LH, Meilahn EN, Plantinga PL** 1991 Weight gain at the time of menopause. Arch Intern Med 151: 97–102
42. **Lovejoy JC, Champagne CM, de Jonge L, Xie H, Smith SR** 2008 Increased visceral fat and decreased energy expenditure during the menopausal transition. Int J Obes (Lond) 32:949–958
43. **Lee CG, Carr MC, Murdoch SJ, Mitchell E, Woods NF, Wener MH, Chandler WL, Boyko EJ, Brunzell JD** 2009 Adipokines, inflammation, and visceral adiposity across the menopausal transition: a prospective study. J Clin Endocrinol Metab 94:1104–1110

OPEN ACCESS Freely available online

PLoS MEDICINE

Medical Evidence of Human Rights Violations against Non-Arabic-Speaking Civilians in Darfur: A Cross-Sectional Study

Alexander C. Tsai[1,2☯]*, Mohammed A. Eisa[3,4☯], Sondra S. Crosby[3,5,6], Susannah Sirkin[3], Michele Heisler[3,7,8,9], Jennifer Leaning[10], Vincent Iacopino[3,11,12]

1 Robert Wood Johnson Health and Society Scholars Program, Harvard University, Cambridge, Massachusetts, United States of America, **2** Center for Global Health, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **3** Physicians for Human Rights, Cambridge, Massachusetts, United States of America, **4** Harvard Humanitarian Initiative, Harvard University, Cambridge, Massachusetts, United States of America, **5** Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, United States of America, **6** Department of Health Law, Bioethics, and Human Rights, Boston University School of Public Health, Boston, Massachusetts, United States of America, **7** Veterans Affairs Center for Clinical Management Research, Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, Michigan, United States of America, **8** Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, United States of America, **9** Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, Michigan, United States of America, **10** Francois-Xavier Bagnoud Center for Health and Human Rights, Harvard School of Public Health, Boston, Massachusetts, United States of America, **11** Department of Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, United States of America, **12** Human Rights Center, University of California at Berkeley, Berkeley, California, United States of America

Abstract

Background: Ongoing conflict in the Darfur region of Sudan has resulted in a severe humanitarian crisis. We sought to characterize the nature and geographic scope of allegations of human rights violations perpetrated against civilians in Darfur and to evaluate their consistency with medical examinations documented in patients' medical records.

Methods and Findings: This was a retrospective review and analysis of medical records from all 325 patients seen for treatment from September 28, 2004, through December 31, 2006, at the Nyala-based Amel Centre for Treatment and Rehabilitation of Victims of Torture, the only dedicated local provider of free clinical and legal services to civilian victims of torture and other human rights violations in Darfur during this time period. Among 325 medical records identified and examined, 292 (89.8%) patients from 12 different non-Arabic-speaking tribes disclosed in the medical notes that they had been attacked by Government of Sudan (GoS) and/or Janjaweed forces. Attacks were reported in 23 different rural council areas throughout Darfur. Nearly all attacks (321 [98.8%]) were described as having occurred in the absence of active armed conflict between Janjaweed/GoS forces and rebel groups. The most common alleged abuses were beatings (161 [49.5%]), gunshot wounds (140 [43.1%]), destruction or theft of property (121 [37.2%]), involuntary detainment (97 [29.9%]), and being bound (64 [19.7%]). Approximately one-half (36 [49.3%]) of all women disclosed that they had been sexually assaulted, and one-half of sexual assaults were described as having occurred in close proximity to a camp for internally displaced persons. Among the 198 (60.9%) medical records that contained sufficient detail to enable the forensic medical reviewers to render an informed judgment, the signs and symptoms in all of the medical records were assessed to be consistent with, highly consistent with, or virtually diagnostic of the alleged abuses.

Conclusions: Allegations of widespread and sustained torture and other human rights violations by GoS and/or Janjaweed forces against non-Arabic-speaking civilians were corroborated by medical forensic review of medical records of patients seen at a local non-governmental provider of free clinical and legal services in Darfur. Limitations of this study were that patients seen in this clinic may not have been a representative sample of persons alleging abuse by Janjaweed/GoS forces, and that most delayed presenting for care. The quality of documentation was similar to that available in other conflict/post-conflict, resource-limited settings.

Please see later in the article for the Editors' Summary.

Citation: Tsai AC, Eisa MA, Crosby SS, Sirkin S, Heisler M, et al. (2012) Medical Evidence of Human Rights Violations against Non-Arabic-Speaking Civilians in Darfur: A Cross-Sectional Study. PLoS Med 9(4): e1001198. doi:10.1371/journal.pmed.1001198

Academic Editor: Egbert Sondorp, London School of Hygiene and Tropical Medicine, United Kingdom

Received July 15, 2011; **Accepted** February 16, 2012; **Published** April 3, 2012

Funding: This study was funded by Physicians for Human Rights with support from the Open Society Foundations and the Sudan Aid Fund of the Community Foundation of Western Massachusetts. ACT received salary support through the Robert Wood Johnson Health & Society Scholars Program. MAE received salary support through Physicians for Human Rights, the Open Society Foundations, the Sudan Aid Fund of the Community Foundation of Western Massachusetts, the Scholar Rescue Fund of the Institute for International Education, and the Robert F. Kennedy Center for Justice and Human Rights, as well as in-kind support from the Harvard Humanitarian Initiative. No funding bodies had any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Six of the study authors are currently affiliated with organizations that advocate for the prevention of mass atrocities, protection of internationally guaranteed rights, and/or prosecution of those who violate human rights: the Amel Centre for Treatment and Rehabilitation of Victims of Torture (MAE), Physicians for Human Rights (MAE, SC, SS, MH, VI), the Harvard Humanitarian Initiative (MAE), the Francois-Xavier Bagnoud Center for Health and Human Rights at the Harvard School of Public Health (JL), and the Human Rights Center at the University of California at Berkeley (VI). From 2004 to 2009, one of the study authors (MAE) served as medical director of the clinic from which the data were obtained. All authors have declared that no financial conflicts of interest exist.

Abbreviations: ICC, International Criminal Court; IDP, internally displaced person; GoS, Government of Sudan

* E-mail: atsai@hsph.harvard.edu

☯ These authors contributed equally to this work.

Introduction

In the Darfur region of western Sudan, ongoing conflict between Arabic-speaking and non-Arabic-speaking tribes [1,2] has reached crisis proportions since the Government of Sudan (GoS) first initiated its military response to organized armed groups opposing the GoS [2]. In addition to targeting armed rebel forces in its response, however, the GoS has also been accused of targeting non-Arabic-speaking civilians, namely members of the Fur, Masalit, and Zaghawa tribes [3,4]. By the end of 2007, more than 2.4 million refugees from the violence, or nearly one-third of the population [5], had fled to camps for internally displaced persons (IDPs) within Darfur or to similar refugee camps in neighboring Chad [6], thus creating a severe humanitarian disaster.

Prior research has focused on generating accurate mortality estimates to inform policy and programming [7–13], with recent studies estimating 200,000–300,000 deaths directly and indirectly attributable to the conflict in 2003–2005 alone [14,15]. The reported systematic, repeated, targeted assaults on non-Arabic-speaking civilians, large-scale disruption of rural livelihoods, and deliberate consignment to conditions conducive to death have prompted observations that these could constitute acts of genocide [4,5,16–19]. Following a United Nations–appointed Commission of Inquiry and an International Criminal Court (ICC) investigation, the Pre-Trial Chamber I of the ICC issued arrest warrants for allegedly responsible authorities, including two arrest warrants for Sudanese President Omar Hassan Ahmad Al Bashir (''Al Bashir'') on the grounds of crimes against humanity (March 4, 2009) [20,21] and genocide (July 12, 2010) [22].

Despite investigations into the violence in Darfur, little research to date has been able to make use of Sudanese documents to substantiate victims' or observers' claims of violence amounting to war crimes, crimes against humanity, or genocide. GoS forces were implicated in the Atrocities Documentation Survey [5,23–25]. Arab Janjaweed (''men with guns on horses or camels'') militias, which originated as Libyan proxy militias in the Chadian civil war and have been suspected of collaborating with GoS forces [2], have been implicated in reports of sexual violence described by Darfuri women now living in IDP camps [26]. The systematic destruction of livelihoods, which under certain circumstances can be considered an act of genocide [17,27,28], has also been described. However, a critical limitation of prior studies is their reliance on self-report data gathered from victims living in refugee camps outside of Sudan. One team of investigators attempted to conduct interviews at IDP camps within Darfur but was refused access by the GoS [17,27]. With unique access to medical records of clinical encounters in Darfur, we undertook this study to characterize the nature and geographic scope of alleged abuses against civilians in Darfur and to substantiate the allegations with forensic review and analysis of the medical evidence.

Methods

Ethics Statement

As this was a retrospective analysis of de-identified medical records, informed consent was not obtained. All study procedures were approved by the Harvard School of Public Health Office of Human Research Administration as well as an independent ethics review board convened for this research project by Physicians for Human Rights. Given that the medical records used in the analysis were de-identified, this research project was assessed to represent minimal risk to Amel Centre patients. The ethics review board was guided by the relevant process provisions of Title 45 of the US Code of Federal Regulation and the Declaration of Helsinki as revised in 2000 [29] and was composed of individuals with expertise in forensic medicine, public health, bioethics, and international health and human rights research.

Study Population and Setting

Data for this study consisted of 325 de-identified medical records of all initial visits (i.e., excluding follow-up visits) of patients seen for treatment at the Amel Centre for Treatment and Rehabilitation of Victims of Torture, in Nyala, South Darfur, from its opening on September 28, 2004, through December 31, 2006. Records for 2007–2009 could not be retrieved because of ongoing security concerns (as described in more detail below). With funding from the European Commission, the United Nations High Commissioner for Refugees, the US Agency for International Development, and the United Nations Development Programme, the Amel Centre was the only dedicated local non-governmental provider of free clinical and legal services to any civilian victim of torture or other human rights violations. The Amel Centre received referrals from volunteers placed in three large IDP camps near Nyala (Dreig, Otash, and Kalma) but accepted civilian clients from all over Darfur. Aside from the free services, and transportation to and from the IDP camps, patients were not given additional incentives or benefits.

The Amel Centre's initial staff in the Nyala office consisted of one general medicine doctor (M. A. E.), one junior doctor, one psychosocial worker, and two lawyers. Their training on the treatment of victims of torture and sexual violence was facilitated by the Sudan Organisation Against Torture and was consistent with the *Manual on Effective Investigation and Documentation of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment* (''Istanbul Protocol'') [30–32]. The paper-based record-keeping system was similar to other prototypical clinics operating in conflict zones. Although examining clinicians typically conversed with patients in the patients' language of choice (typically Fur, Zaghawa, or Dago), notes documenting the encounters were written in English. A standardized medical record form was used, but few fields specified closed-ended responses (e.g., ''name,'' ''date of birth,'' ''date of detention''). The content of the clinical encounter, and therefore the bulk of the medical record documentation, was driven by patients' concerns. A network of volunteer physicians and social workers provided specialty care, and all women who disclosed that they had been sexually assaulted were referred to a gynecologist for evaluation. The laboratories associated with the network were able to implement all necessary blood, urine, stool, serological, and pregnancy tests but did not have the capacity for deoxyribonucleic acid analysis. After the initial visit, follow-up visits were provided to assess symptomatic improvement and provide longer term physiotherapy where indicated. The care provided and coordinated by the Amel Centre was delivered under difficult and often dangerous field conditions. After the ICC issued the first arrest warrant for Al Bashir [21], the Sudanese Ministry of Humanitarian Affairs ordered the Amel Centre, along with two other local and 13 foreign aid groups, to cease operations [33]. The three clinicians formerly on staff are now living abroad. Prior to fleeing the country, they preserved hard copies of the medical records for 2004–2006 and sent them out of the country.

Data Collection

Amel Centre clinicians generated the medical records for the purposes of clinical care and internal record-keeping. We sought

to abstract the data both accurately and in such a way as to capture the main themes identified in the records. Guided by prior research [34–41], we created a list of different types of abuses that may be considered evidence of torture and/or other human rights violations, as well as symptoms potentially reported by patients and signs potentially documented by examining clinicians. Then we developed a medical record abstraction tool that included lists of standardized names (e.g., tribes and rural council areas) and response options to guide efficient abstraction of data (Text S1).

Data were abstracted by one of the authors (A. C. T.) from the demographic, incident, and clinical care components of the Amel Centre's general medical records. Although the general medical records may have noted the presence of diagnostic or laboratory testing, or specialty medical records, access to these raw data elements unfiltered by the examining clinicians (e.g., diagnostic imaging or laboratory reports) were not available for our analysis. We collected data on patient socio-demographic characteristics, alleged abuses, and the harms reportedly resulting from the abuses. To assess the reliability of the data abstraction tool for coding the key variables on alleged abuses, we randomly selected 20 medical records for independent coding by two other study authors (M. A. E. and S. S. C.) as well as for wider discussion by the research team. Inter-rater agreement exceeded 0.70 on the coding of most of these variables. The abstraction tool was further refined through an iterative process to ensure that the variables were clearly defined and could be applied consistently to the data. With regards to content validity, the 20 records were carefully reviewed to ensure that all potential categories were represented. Data from the remaining 305 medical records were then abstracted by a single author (A. C. T.).

Medical Opinions on Alleged Abuses

Two study authors with extensive medical experience in the evaluation and treatment of survivors of torture and other forms of physical and psychological abuse (S. S. C. and V. I.) independently reviewed each medical record, blinded to each other's assessments. First, they determined whether the medical record contained sufficient detail to enable an informed opinion about the consistency of the documented signs and symptoms with the record of alleged abuses in the medical notes. Second, among the cases that did contain sufficient detail, they assessed the extent to which the recorded signs and symptoms were consistent with the alleged abuses described in the medical record. Consistency was scored using a five-point Likert-type scale: "not related to alleged abuse," "not consistent with," "consistent with," "highly consistent with," and "virtually diagnostic of." These evaluations were based on the Istanbul Protocol [30–32] and other conventions for the evaluation of survivors of torture and other human rights abuses [42–45].

Statistical Analysis

Data were entered into Excel (version 12.0, Microsoft) and then exported to Stata (version 11.0, StataCorp) for analysis. We characterized socio-demographic, incident, and clinical variables with medians and inter-quartile ranges. Inter-rater agreement was assessed using the kappa statistic [46]. The locations of alleged attacks were mapped to the longitude and latitude [47–49] of the administrative center, principal town, or largest secondary town of the rural council area where they were reported to have taken place. ArcGIS (version 9.2, Esri) was used to generate a continuously variable proportional circle map, with circle sizes classified into seven categories using the Fisher-Jenks algorithm [50].

Results

Characteristics of Amel Centre Patients

We obtained medical records for all 325 patients presenting for care at the Amel Centre from September 28, 2004, to December 31, 2006. Summary statistics are presented in Table 1. Most patients were brought in by friends or relatives (54.2%) or by staff or volunteers (28.0%). Median age was 35 y, with a range of 4–82 y. Thirty patients (9.2%) were under the age of 18 y. Men comprised the majority of patients (252 [77.5%]). Approximately one-half (49.5%) of the men and boys were younger than 36 y of age. Most patients were married (76.3%). All self-identified as Muslim. The sample included patients from 14 different non-Arabic-speaking tribes, and members of the Fur, Zaghawa, and Dago tribes accounted for nearly 90% of patients. Only two (0.6%) patients were from Arab tribes. Most (84.9%) lived in South Darfur, where the Amel Centre was located.

Patterns and Geographic Scope of Alleged Abuses

The attacks documented in the patients' records occurred between 2002 and 2006, with a peak frequency in March 2005. Characteristics of these incidents are displayed in Table 2. Between the date of the incident and the date of presentation at the Amel Centre, a median of 101 d had elapsed (inter-quartile range, 22–365 d). Approximately one-third (36.6%) of patients presented to the Amel Centre within 6 wk of the alleged incident.

Alleged attacks on individuals and villages recorded in the patients' records took place in 23 rural council areas (out of 65 total) throughout Darfur (Figure 1). Of the total, 281 (86.5%) occurred in South Darfur, 35 (10.8%) occurred in West Darfur, and eight (2.5%) occurred in North Darfur. Approximately one-third (35.1%) of the attacks disclosed by patients were also described by at least one other Amel Centre patient. Many villages were repeatedly attacked, with five villages reportedly attacked a total of 41 times during the study period: Marla was attacked 13 times during a 12-mo period from December 2004 to December 2005; Adwa, ten times (October 2003–November 2005); Labado, seven times (March 2004–December 2005); Bendisi, six times (August 2003–Dececember 2004); and Mukjar, five times (August 2003–August 2004). In addition, 46 (14.2%) patients disclosed that they had been attacked in the vicinity of an IDP camp: 16 (34.8%) of these attacks occurred inside the camp, 15 (32.6%) occurred a median of 3 km outside the camp, six (13.0%) occurred an unspecified distance outside the camp, and nine (19.6%) occurred within the general vicinity of a camp but the exact location was unspecified.

Two hundred ninety-three (90.1%) patients described their perpetrators as GoS and/or Janjaweed forces; of these, 48 (16.4%) stated that GoS and Janjaweed forces attacked in concert (Table 3). Thirty-two (9.9%) patients disclosed that they had been attacked by rebel soldiers, bandits, community authorities, or other community members. Among those attacked by GoS and/or Janjaweed, 292 (99.7%) patients were from 12 different non-Arabic-speaking tribes, and only one (0.3%) was from an Arab tribe. Thirty-two (9.9%) patients disclosed to the examining clinician that a military commander was present during the attack. Nearly all (98.8%) attacks occurred in the absence of active armed conflict between GoS/Janjaweed forces and rebel groups. The examining clinician noted when patients speculated as to reasons for the attack: 60 (18.5%) patients stated that they were targeted because the attackers suspected them of being rebels, and 58 (17.9%) stated that they were targeted because of their racial or tribal identity.

Table 1. Characteristics of patients presenting for care at the Amel Centre for Treatment and Rehabilitation of Victims of Torture in Nyala, South Darfur.

Variable Name	Number (Percent)
Referral source	
Brought to center by friend/relative	176 (54.2%)
Brought to center by staff/volunteer	91 (28.0%)
Self-referral	44 (13.5%)
Referred by friend/relative	13 (4.0%)
Visit year	
2004	47 (14.5%)
2005	233 (71.7%)
2006	45 (13.9%)
Sex	
Male	252 (77.5%)
Female	73 (22.5%)
Age	
<26 y	96 (29.5%)
26–35 y	85 (26.2%)
36–45 y	68 (20.9%)
46–55 y	43 (13.2%)
>55 y	32 (9.9%)
Marital status	
Single	77 (23.7%)
Married	248 (76.3%)
Has children	
Yes	173 (53.2%)
No	3 (0.9%)
Unknown/unspecified	149 (45.9%)
Religion	
Muslim	325 (100%)
Other	0
Tribe	
Fur	173 (53.2%)
Zaghawa	76 (23.4%)
Dago	38 (11.7%)
Bargo	7 (2.2%)
Other	31 (9.5%)
Occupation/profession	
Farmer	199 (61.2%)
Student	42 (12.9%)
Merchant	26 (8.0%)
Unemployed	12 (3.7%)
Other	46 (14.2%)
State of residence	
South Darfur	276 (84.9%)
West Darfur	37 (11.4%)
North Darfur	5 (1.5%)
Unknown/unspecified	7 (2.2%)

doi:10.1371/journal.pmed.1001198.t001

Table 2. Characteristics of incidents that led to injuries.

Variable Name	Number (Percent) or Median (IQR)
Incident year	
2002	4 (1.2%)
2003	43 (13.2%)
2004	162 (50.0%)
2005	89 (27.4%)
2006	26 (8.0%)
Unknown/unspecified	1 (0.3%)
Days elapsed between incident and presentation to Amel Centre	101 (22–365)
Days elapsed ≤42 d	
Yes	119 (36.6%)
No	204 (62.8%)
Could not be calculated	2 (0.6%)
Same incident also reported by another Amel Centre patient	
Yes	114 (35.1%)
Unspecified	211 (64.9%)
Rural council area where incident took place	
Nyala	74 (22.8%)
Malam	59 (18.2%)
Abu Agura	34 (10.5%)
Yasin	32 (9.9%)
Shearia	32 (9.9%)
Other locations throughout North, South, and West Darfur	94 (28.9%)
Unknown	1 (0.3%)
IDP camp where incident took place, if noted	
Dreig	17 (37.0%)
Otash	15 (32.6%)
Kalma	14 (30.4%)
Distance from IDP camp	
Outside camp	21 (45.7%)
Inside camp	16 (34.8%)
In the general vicinity (but exact distance unspecified)	9 (19.6%)
Distance outside IDP camp (kilometers)[a]	3 (3–3)

[a]From the 15 records in which the patient provided an estimated distance to the examining clinician.
IQR, inter-quartile range.
doi:10.1371/journal.pmed.1001198.t002

Patients' medical records described a wide range of alleged abuses, including beatings (161 [49.5%]), gunshot wounds (140 [43.1%]), destruction or theft of private property (121 [37.2%]), involuntary detainment (97 [29.9%]), and being bound with rope, chains, or other material (64 [19.7%]) (Table 4). GoS forces were described as accounting for more than one-half of custody-related incidents (61 [59.8%]), whereas Janjaweed forces were alleged to have accounted for most incidents involving physical assault (148 [50.7%]), sexual assault (28 [62.2%]), and destruction or theft of private property (77 [63.6%]). In addition to the abuses patients personally experienced, the medical records for this group of

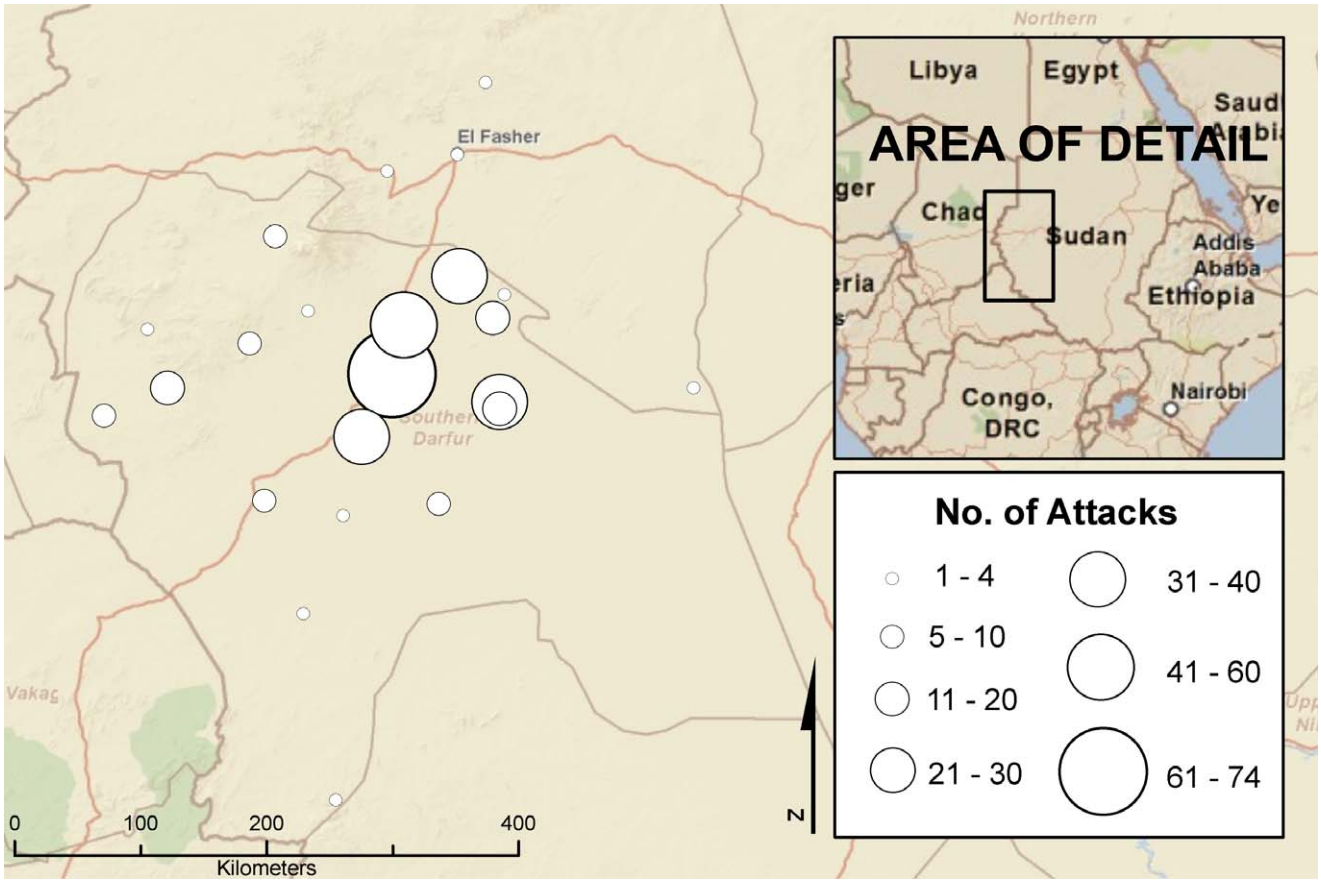


Figure 1. Geographic pattern of attacks reported by patients, 2002–2006. The largest circle corresponds to the town of Nyala, where the Amel Centre was located. The base map for this figure was obtained from ArcGIS (version 9.2, Esri) Online World StreetMap, accessed on February 22, 2011. Sources: Esri, DeLorme, NAVTEQ, TomTom, US Geological Survey, Intermap, Increment P Corporation, Natural Resources Canada, Esri Japan, and the Japanese Ministry of Economy, Trade and Industry.
doi:10.1371/journal.pmed.1001198.g001

patients also collectively describe that they witnessed the killing of 948 other persons.

Consistency between Allegations of Abuse and the Signs and Symptoms Described in the Medical Records

The signs and symptoms most frequently documented in the medical records were chronic pain (194 [59.7%]), wounds or scars (167 [51.4%]), functional disabilities (e.g., contractures causing restricted grasp) (65 [20.0%]), and bone fractures (55 [16.9%]) (Table 5). There was 96.3% agreement ($\kappa = 0.92$) between the medical reviewers on whether the medical records contained sufficient detail to enable an informed opinion about the consistency of the recorded signs and symptoms with the allegations documented in the medical record. More than one-third (127 [39.1%]) of the medical records were assessed by at least one reviewer to lack sufficient detail (i.e., documentation was incomplete) to enable him or her to render an informed judgment about consistency. Of the 198 (60.9%) records that were considered sufficiently detailed by both reviewers, the medical reviewers agreed that the recorded signs and symptoms were either consistent with (101 [51.0%]), highly consistent with (81 [40.9%]), or virtually diagnostic of (5 [2.5%]) the alleged abuses. There were no cases in which the reports of medical examinations were considered not consistent with, or unrelated to, the recorded allegations. In only 11 (3.4%) cases did the medical reviewers disagree in their consistency scorings, for an excellent inter-rater agreement overall ($\kappa = 0.89$).

Approximately one-half (36 [49.3%]) of all women presenting to the Amel Centre disclosed that they had been sexually assaulted. One-half of sexual assaults on women were recorded as having occurred in close proximity to an IDP camp, with nine (25.0%) recorded as having occurred in the general vicinity of the camp and nine (25.0%) having occurred within 3 km of the camp. The majority (31 [86.1%]) of sexual assaults on women involved rape or gang rape. Among these, five (16.1%) women disclosed they had become pregnant as a result of the alleged rape; no follow-up information was available on the remainder. Nine men also disclosed that they had been sexually assaulted, including one who had been raped. Twenty-five (55.6%) medical records of sexual assault victims were considered by the medical reviewers to be sufficiently detailed in recorded signs and symptoms to enable him or her to render an informed judgment about consistency. Of these, the medical reviewers agreed that the medical evidence was consistent with (14 [56.0%]), highly consistent with (9 [36.0%]), or virtually diagnostic of (1 [4.0%]) the alleged sexual assault. There were no cases in which the medical findings were considered not consistent with, or unrelated to, the alleged sexual assault. The reviewers disagreed about the scoring for one (4.0%) case, for an excellent inter-rater agreement on sexual assault findings overall ($\kappa = 0.92$).

Table 3. Characteristics of alleged perpetrators.

Variable Name	Number (Percent) or Median (IQR)
Affiliation of alleged direct perpetrator(s)	
Janjaweed	166 (51.1%)
GoS	79 (24.3%)
Both GoS and Janjaweed	48 (14.8%)
Other	32 (9.9%)
Military commander present	
Yes	32 (9.9%)
Unspecified	293 (90.2%)
Number of alleged perpetrators	
Single perpetrator	7 (2.2%)
More than one but exact number unspecified	207 (63.7%)
2–5 perpetrators	55 (16.9%)
6–10 perpetrators	21 (6.5%)
More than ten perpetrators	35 (10.8%)
Number of alleged perpetrators, if noted	5 (2–20)
Reason for incident as perceived by patient[a]	
Suspected of being a rebel	60 (18.5%)
Targeted because of racial or tribal identity	58 (17.9%)
Suspected of supporting rebels	24 (7.4%)
Suspected of theft, or was defending self against theft	11 (3.4%)
Suspected of political activity	5 (1.5%)
Suspected of working against rebels	4 (1.2%)

[a]Responses in this category were not mutually exclusive, so percentages do not add up to 100.
IQR, inter-quartile range.
doi:10.1371/journal.pmed.1001198.t003

Discussion

We analyzed the medical records of 325 consecutive patients who were seen for care at the Amel Centre in Nyala, Darfur, between September 28, 2004, and December 31, 2006, with the aim of assessing the consistency between the recorded allegations of abuse and the signs and symptoms noted in the medical record. Our findings show that in all of the medical records that contained sufficient detail, the medical evidence was considered to be at least consistent with (if not highly consistent with or virtually diagnostic of) the human rights violations disclosed by the patients. There were no cases in which the reports of medical examinations were considered not consistent with, or unrelated to, the recorded allegations. Most of the abuses described in the medical records—which included beatings, killings, sexual assault, torture, and involuntary detainment—were allegedly perpetrated by GoS and Janjaweed forces and were described as having occurred throughout Darfur, with five villages attacked a total of 41 times during the study period. The spatial distribution of reported incident locations in our data suggests, at a minimum, that the attacks were widespread. However, our lack of a representative population-based sample makes it difficult for us to generalize about the full extent or population incidence of attacks. Many patients reported attacks by GoS and Janjaweed forces acting in concert. In some cases, patients disclosed to the examining clinician the names of specific victims, perpetrators, or military commanders, and this information was noted in the medical record. Fewer than 1% of patients reported observing the perpetrators to be in active armed conflict with rebel or other groups. Although the medical reviewers had no way to corroborate the identities of the perpetrators, these findings are consistent with prior research implicating GoS forces in the perpetration of human rights violations upon non-Arabic-speaking civilians in Darfur [24,25].

Nearly one-half of women presenting for care disclosed that they had been sexually assaulted. The use of sexual violence in armed conflict has been recognized as a means of not only demoralizing individual victims but also destabilizing their families and terrorizing communities [51–56]. Rape and other forms of sexual violence have been recognized as war crimes and crimes against humanity [57], as well as instruments of genocide [55,58]. Moreover, one-half of these assaults were described as having occurred in close proximity to an IDP camp. These data are consistent with prior reports of rapes occurring near IDP camps [26,55,59], as well as previous work documenting that violence was responsible for a substantial proportion of deaths among persons settled (i.e., not in the villages or in flight) in IDP camps in West Darfur [11]. Collectively, these data raise questions about the security provided to persons living in IDP camps, notably women, who must frequently venture outside the camp to gather firewood for fuel [27]. The Inter-Agency Standing Committee has issued guidelines that suggest several minimum prevention and response interventions that could be implemented with regards to security mechanisms instituted in areas of close proximity to IDP camps.

In contrast to prior studies' reliance on self-report of refugees living outside of Darfur [17,24,25,27,28], our data are based on unusual access to medical records of clinical encounters in Darfur maintained by local clinicians directly responsible for treatment and record-keeping. Medical forensic experts reviewed and analyzed the signs and symptoms described in the medical records and evaluated their consistency with the alleged abuses documented in the medical notes. Less than two-thirds of the records were detailed enough for the forensic reviewers to substantiate the patients' claims of abuse, a finding that is not surprising given that the Amel Centre medical records were not intended for research purposes. In a similar study in which third-party experts assessed the official medical evaluations of forensic experts working for the Mexican Procuraduría General de la República (Office of the Attorney General), in 18 of 39 cases (46%) their assessments were indeterminate because of insufficient documentation to corroborate alleged torture and ill treatment [60]. Their findings are consistent with ours and highlight the potential value of using clinical information to corroborate allegations of abuse. In our study, among the medical records that contained sufficient detail, all were assessed to be at least consistent with (if not highly consistent with or virtually diagnostic of) the allegations. These data substantially enhance the credibility of the patients' claims of abuse. Importantly, however, the medical records provided the forensic reviewers with no data that could be used to corroborate either claims of assailant identities or claims of genocidal intent.

Limitations

Interpretation of our findings is subject to a number of limitations. First, we used a discrete, comprehensive sample of patients, but it was not systematic. During this time period, the Sudanese Criminal Procedure Act required all injury or trauma victims to file a report with the police in order to obtain a medical evidence form ("Form 8"), without which they were legally not permitted to receive treatment from an authorized medical officer [55,61–63]. In practice the police were known to deny the Form 8

Table 4. Types of abuses disclosed by patients.

Type of Abuse	Affiliation of Alleged Perpetrator(s)	
	GoS and/or Janjaweed ($n=293$)	Other/Unknown ($n=32$)
Attacks involving heavy weapons	33 (11.3%)	2 (6.3%)
Ground explosives (bombing, grenades)	2 (0.7%)	
Attack by aircraft or helicopter	18 (6.1%)	1 (3.1%)
Attack by land cruiser	24 (8.2%)	1 (3.1%)
Physical assault	264 (90.1%)	28 (87.5%)
Blunt trauma (beating, whipping)	145 (49.5%)	16 (50.0%)
Gunshot wound	125 (42.7%)	15 (46.9%)
Burns or electric shocks	21 (7.2%)	
Stretch injury (hanging, suspension)	19 (6.5%)	
Genital trauma	10 (3.4%)	
Sexual assault	39 (13.3%)	6 (18.8%)
Forced undressing	12 (4.1%)	1 (3.1%)
Insertion of foreign object into anus/vagina	3 (1.0%)	
Attempted rape	5 (1.7%)	
Rape	15 (5.1%)	1 (3.1%)
Rape by more than a single perpetrator	12 (4.1%)	4 (12.5%)
Humiliation or psychological manipulation	70 (23.9%)	3 (9.4%)
Verbal abuse	32 (10.9%)	1 (3.1%)
Verbal abuse involving racial slurs	6 (2.1%)	
Forced performance of humiliating/taboo acts	7 (2.4%)	
Verbalized threats of death	43 (14.7%)	2 (6.3%)
Custody-related violations	95 (32.4%)	7 (21.9%)
Involuntary detainment	90 (30.7%)	7 (21.9%)
Bound with rope or other apparatus	60 (20.5%)	4 (12.5%)
Crowded, unhygienic conditions	43 (14.7%)	
Deprived of food/water or medical care	32 (10.9%)	1 (3.1%)
Sensory deprivation	25 (8.5%)	
Destruction or theft of private property	115 (39.3%)	6 (18.8%)

Data are number (percent).
doi:10.1371/journal.pmed.1001198.t004

Table 5. Common symptoms and signs documented in patient medical records.

Type of Symptom or Sign	Number (Percent)
Pain (non-pelvic)	194 (59.7%)
Wounds or scars	167 (51.4%)
Functional disability	65 (20.0%)
Broken or fractured bones	55 (16.9%)
Weakness	38 (11.7%)
Pelvic pain	34 (10.5%)
Insomnia	32 (9.9%)
Numbness	23 (7.1%)
Swelling	18 (5.5%)
Headache	14 (4.3%)

doi:10.1371/journal.pmed.1001198.t005

to members of non-Arabic-speaking tribes, so this legal requirement represented a substantial hurdle, and in many cases a complete barrier, to accessing health care services. Consistent with this, patients in our sample presented for care a median of 101 d after the abuses leading to their need for treatment. Furthermore, the majority of patients seen were men, highlighting the issue of women's lack of adequate access to care and their overall limited public mobility in this setting. These barriers are of particular salience with regards to cases of sexual violence [61], where victims may fear reprisals, blame, and other psychosocial consequences of disclosure [5,26,59,64–68]. A second limitation relates to the delay in presentation for care. Although physical and psychological sequelae may persist for years and even for the duration of a victim's lifetime, some symptoms and disabilities may resolve or diminish over time [42–45,69]. Despite their training and experience, Amel Centre staff could have under-detected and therefore under-documented some symptoms, especially those concerning sensitive topics (e.g., sexual violence, psychological distress) that might not be spontaneously disclosed. More generally, the medical chart review literature is characterized by under-documentation of signs and symptoms [70–73], so we

would expect this limitation to generically apply in any setting. Third, because these data were not collected in a research setting, in most if not all cases, the same individual documented both the allegations of abuse and the results of the medical examination. The examining clinician's prior knowledge of the nature of the allegations could have biased the completeness of the documentation with regards to signs and symptoms observed. Fourth, few rape cases were scored by our medical forensic experts as virtually diagnostic of the alleged assault. Amel Centre protocol directed all female rape victims to a gynecologist for evaluation [74]. However, these records were unavailable for analysis because they could not be secured and sent out of the country prior to the clinicians' fleeing the country (as described above). Fifth, we were unable to include information on victims who were killed, so it may be more appropriate to regard our data as underestimating the true severity of atrocities inflicted upon non-Arabic-speaking civilians living in this region. Sixth, Amel Centre staff were routinely subject to surveillance, detainment, and interrogation by GoS forces [75,76]. With increasing frequency in 2009, Amel Centre staff were detained, interrogated, tortured, and accused of collaborating with the ICC. Upon Al Bashir's indictment, they were advised to flee the country. Because of ongoing security concerns, we could not obtain the records for 2007–2009 to analyze for this study. This limitation underscores that the Amel Centre clinicians provided medical and legal services under dangerous working conditions. Health care workers in other settings have faced similar challenges [77], further emphasizing the need for international support for the protection of health professionals working under similar circumstances.

In summary, despite these unavoidable limitations, our study of non-Arabic-speaking civilian patients who visited the Amel Centre in Nyala, Darfur, between 2004 and 2006 found that in all of the medical records that contained sufficient detail, the recorded medical evidence was considered at least consistent with the alleged incidents of torture and other human rights violations. There were no cases in which the reports of medical examinations were considered not consistent with, or unrelated to, the recorded allegations. The widespread, organized, and sustained pattern of attacks documented in our study indicates that the actions of Janjaweed and GoS forces may constitute war crimes, crimes against humanity, and/or possibly acts of genocide.

Supporting Information

Text S1 Coding sheet, with lists of standardized names and response options, used to guide abstraction of data from the medical records (version of December 7, 2010).
(PDF)

Acknowledgments

We thank the patients, for coming forward to seek care and share their experiences with the Amel Centre staff; the Amel Centre staff, for their courage and commitment to the health and well-being of the people of Darfur despite extreme conditions and great personal risk; Holly Atkinson, M.D., Adrienne Fricke, J.D., M.A., Frank Davidoff, M.D., and Mark Russell, M.A., for critically reading a draft of the manuscript and providing helpful comments; and Jeff Blossom, M.A., of the Center for Geographic Analysis at Harvard University, for his assistance in mapping the attacks to specified geographic locations.

Author Contributions

Conceived and designed the experiments: ACT MAE SSC SS MH JL VI. Analyzed the data: ACT. Wrote the first draft of the manuscript: ACT. Contributed to the writing of the manuscript: ACT MAE SSC VI. ICMJE criteria for authorship read and met: ACT MAE SSC SS MH JL VI. Agree with manuscript results and conclusions: ACT MAE SSC SS MH JL VI. Obtained funding for the study: MAE SS JL. Acquired the data: ACT MAE SS VI. Contributed to interpretation of the data: ACT MAE SSC SS MH JL VI.

References

1. Mullins CW, Rothe DL (2007) The forgotten ones: the Darfuri genocide. Crit Criminol 15: 135–158.
2. Flint J, De Waal A (2005) Darfur: a short history of a long war. London: Zed Books.
3. Hertzke AD (2005) The shame of Darfur. First Things 156: 16–22.
4. Straus S (2005) Darfur and the genocide debate. Foreign Aff 84: 123–133.
5. US Department of State (2004) Documenting atrocities in Darfur. State publication 11182. Washington (District of Columbia): Bureau of Democracy, Human Rights, and Labor and the Bureau of Intelligence and Research, US Department of State.
6. Office of UN Deputy Special Representative of the UN Secretary General for Sudan, UN Resident and Humanitarian Co-ordinator (2008) Darfur humanitarian profile no. 30. Khartoum: United Nations.
7. Nabarro D (2004) Mortality projections for Darfur. Geneva: World Health Organization.
8. Hagan J, Rymond-Richmond W, Parker P (2005) The criminology of genocide: the death and rape of Darfur. Criminology 43: 525–561.
9. US Government Accountability Office (2006) Darfur crisis: death estimates demonstrate severity of crisis, but their accuracy and credibility could be enhanced. Washington (District of Columbia): US Government Accountability Office.
10. McCarthy B, Hagan J (2009) Counting the deaths in Darfur: pitfalls on the pathway to a public sociology. In: Jeffries V, ed. Handbook of public sociology. Lanham: Rowman and Littlefield Publishers. pp 319–337.
11. Depoortere E, Checchi F, Broillet F, Gerstl S, Minetti A, et al. (2004) Violence and mortality in West Darfur, Sudan (2003–04): epidemiological evidence from four surveys. Lancet 364: 1315–1320.
12. Grandesso F, Sanderson F, Kruijt J, Koene T, Brown V (2005) Mortality and malnutrition among populations living in South Darfur, Sudan: results of 3 surveys, September 2004. JAMA 293: 1490–1494.
13. Checchi F, Roberts L (2008) Documenting mortality in crises: what keeps us from doing better. PLoS Med 5: e146. doi:10.1371/journal.pmed.0050146.
14. Hagan J, Palloni A (2006) Social science. Death in Darfur. Science 313: 1578–1579.
15. Degomme O, Guha-Sapir D (2010) Patterns of mortality rates in Darfur conflict. Lancet 375: 294–300.
16. Leaning J (2004) Diagnosing genocide—the case of Darfur. N Engl J Med 351: 735–738.
17. VanRooyen MJ, Leaning J, Johnson K, Hirschfeld K, Tuller D, et al. (2008) Employment of a livelihoods analysis to define genocide in the Darfur region of Sudan. J Genocide Res 10: 343–358.
18. Leaning J (2007) The crime of genocide. In: Beyrer C, Pizer HF, eds. Public health and human rights. Baltimore: The Johns Hopkins University Press. pp 202–221.
19. Powell CL (2004) The crisis in Darfur. Testimony before the U.S. Senate Foreign Relations Committee, September 9, 2004. Washington (District of Columbia): US Department of State.
20. Ssenyonjo M (2010) II. The International Criminal Court arrest warrant decision for President Al Bashir of Sudan. Int Comp Law Q 59: 205–225.
21. International Criminal Court (2009) Warrant of arrest for Omar Hassan Ahmad Al Bashir. ICC-02/05-01/09. The Hague: International Criminal Court.
22. International Criminal Court (2010) Second warrant of arrest for Omar Hassan Ahmad Al Bashir. ICC-02/05-01/09. The Hague: International Criminal Court.
23. Totten S, Markusen E (2005) The U.S. government Darfur genocide investigation. J Genocide Res 7: 279–290.
24. Hagan J, Rymond-Richmond W, Palloni A (2009) Racial targeting of sexual violence in Darfur. Am J Public Health 99: 1386–1392.
25. Hagan J, Rymond-Richmond W (2008) The collective dynamics of racial dehumanization and genocidal victimization in Darfur. Am Sociol Rev 73: 875–902.
26. Physicians for Human Rights (2009) Nowhere to turn: failure to protect, support and assure justice for Darfuri women. Cambridge: Physicians for Human Rights and Harvard Humanitarian Initiative.
27. Physicians for Human Rights (2006) Darfur: assault on survival: a call for security, justice and restitution. Cambridge: Physicians for Human Rights.
28. Buchanan-Smith M, Jaspars S (2007) Conflict, camps and coercion: the ongoing livelihoods crisis in Darfur. Disasters 31(Suppl 1): S57–S76.
29. US Department of Health and Human Services (2004) Guidelines for the conduct of research involving human subjects at the National Institutes of Health. Washington (District of Columbia): US Department of Health and Human Services.

30. United Nations High Commissioner on Human Rights (1999) Manual on the effective investigation and documentation of torture and other cruel, inhuman or degrading treatment or punishment. New York: United Nations.
31. Peel M, Iacopino V (2002) The medical documentation of torture. London: Greenwich Medical Media.
32. Iacopino V (2008) Model curriculum on the effective medical documentation of torture and ill-treatment. Copenhagen: International Rehabilitation Council for Torture Victims.
33. Zarocostas J (2009) Expulsion of 13 relief agencies from Darfur could devastate health care. BMJ 338: b985.
34. Basoglu M, Livanou M, Crnobaric C (2007) Torture vs other cruel, inhuman, and degrading treatment: is the distinction real or apparent? Arch Gen Psychiatry 64: 277–285.
35. Sollom R, Richards AK, Parmar P, Mullany LC, Lian SB, et al. (2011) Health and human rights in Chin State, Western Burma: a population-based assessment using multistaged household cluster sampling. PLoS Med 8: e1001007. doi:10.1371/journal.pmed.1001007.
36. Hougen HP (1988) Physical and psychological sequelae to torture. A controlled clinical study of exiled asylum applicants. Forensic Sci Int 39: 5–11.
37. Mollica RF, Wyshak G, Lavelle J (1987) The psychosocial impact of war trauma and torture on Southeast Asian refugees. Am J Psychiatry 144: 1567–1572.
38. Mollica RF, Caspi-Yavin Y, Bollini P, Truong T, Tor S, et al. (1992) The Harvard Trauma Questionnaire. Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. J Nerv Ment Dis 180: 111–116.
39. Hollifield M, Eckert V, Warner TD, Jenkins J, Krakow B, et al. (2005) Development of an inventory for measuring war-related events in refugees. Compr Psychiatry 46: 67–80.
40. Mollica RF, Caspi-Yavin Y (1991) Measuring torture and torture-related symptoms. Psychol Assess 3: 581–587.
41. Iacopino V, Xenakis SN (2011) Neglect of medical evidence of torture in Guantanamo Bay: a case series. PLoS Med 8: e1001027. doi:10.1371/journal.pmed.1001027.
42. Forrest D (1995) The physical after-effects of torture. Forensic Sci Int 76: 77–84.
43. Goldfeld AE, Mollica RF, Pesavento BH, Faraone SV (1988) The physical and psychological sequelae of torture. Symptomatology and diagnosis. JAMA 259: 2725–2729.
44. Boehnlein JK, Kinzie JD (1995) Refugee trauma. Transcult Psychiatry 32: 223–252.
45. Weinstein HM, Dansky L, Iacopino V (1996) Torture and war trauma survivors in primary care practice. West J Med 165: 112–118.
46. Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20: 37–46.
47. Humanitarian Information Centre for Darfur (2005) South Darfur field atlas. Khartoum: Humanitarian Information Centre for Darfur.
48. Humanitarian Information Centre for Darfur (2005) West Darfur field atlas. Khartoum: Humanitarian Information Centre for Darfur.
49. Humanitarian Information Centre for Darfur (2005) North Darfur field atlas. Khartoum: Humanitarian Information Centre for Darfur.
50. Slocum TA (1999) Thematic cartography and visualization. Upper Saddle River: Prentice Hall.
51. Swiss S, Giller JE (1993) Rape as a crime of war. A medical perspective. JAMA 270: 612–615.
52. Amowitz LL, Reis C, Lyons KH, Vann B, Mansaray B, et al. (2002) Prevalence of war-related sexual violence and other human rights abuses among internally displaced persons in Sierra Leone. JAMA 287: 513–521.
53. Kivlahan C, Ewigman N (2010) Rape as a weapon of war in modern conflicts. BMJ 340: c3270.
54. Mullins CW (2009) "He would kill me with his penis": genocidal rape in Rwanda as a state crime. Crit Criminol 17: 15–33.
55. Gingerich T, Leaning J (2004) The use of rape as a weapon of war in the conflict in Darfur, Sudan. Washington (District of Columbia): Office of Transition Initiatives, US Agency for International Development.
56. Kelly J, Kabanga J, Cragin W, Alcayna-Stevens L, Haider S, et al. (2011) 'If your husband doesn't humiliate you, other people won't': gendered attitudes towards sexual violence in eastern Democratic Republic of Congo. Glob Public Health 7: 285–298.
57. United Nations (1998) United Nations: Rome statute of the International Criminal Court [adopted July 17, 1998]. Int Legal Mat 37: 999–1069.
58. Eboe-Osuji C (2007) Rape as genocide: some questions arising. J Genocide Res 9: 251–273.
59. Médecins Sans Frontières (2005) The crushing burden of rape: sexual violence in Darfur. Amsterdam: Médecins Sans Frontières.
60. Moreno A, Iacopino V (2008) Forensic investigations of torture and ill-treatment in Mexico. A follow-up study after the implementation of the Istanbul Protocol. J Leg Med 29: 443–478.
61. Fricke AL, Khair A (2007) Laws without justice: an assessment of Sudanese laws affecting survivors of rape. Washington (District of Columbia): Refugees International.
62. Human Rights Watch (2008) Five years on: no justice for sexual violence in Darfur. New York: Human Rights Watch.
63. United Nations High Commissioner on Human Rights (2005) Access to justice for victims of sexual violence. Geneva: Office of the United Nations High Commissioner for Human Rights.
64. Koss MP (1993) Detecting the scope of rape: a review of prevalence research methods. J Interpers Violence 8: 198–222.
65. Askin KD (2006) Prosecuting gender crimes committed in Darfur: holding leaders accountable for sexual violence. In: Totten S, Markusen E, eds. Genocide in Darfur: investigating the atrocities in the Sudan. New York: Routledge. pp 141–160.
66. Ellsberg M, Heise L, Pena R, Agurto S, Winkvist A (2001) Researching domestic violence against women: methodological and ethical considerations. Stud Fam Plann 32: 1–16.
67. Watts C, Zimmerman C (2002) Violence against women: global scope and magnitude. Lancet 359: 1232–1237.
68. Amnesty International (2004) Sudan, Darfur: rape as a weapon of war—sexual violence and its consequences. London: Amnesty International.
69. Moreno A, Grodin MA (2000) Photo/essay: the not-so-silent marks of torture. JAMA 284: 538.
70. McDermott MF, Lenhardt RO, Catrambone CD, Walter J, Weiss KB (2006) Adequacy of medical chart review to characterize emergency care for asthma: findings from the Illinois Emergency Department Asthma Collaborative. Acad Emerg Med 13: 345–348.
71. Luck J, Peabody JW, Dresselhaus TR, Lee M, Glassman P (2000) How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am J Med 108: 642–649.
72. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M (2000) Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. JAMA 283: 1715–1722.
73. Dresselhaus TR, Peabody JW, Lee M, Wang MM, Luck J (2000) Measuring compliance with preventive care guidelines: standardized patients, clinical vignettes, and the medical record. J Gen Intern Med 15: 782–788.
74. Jewkes R, Christofides N, Vetten L, Jina R, Sigsworth R, et al. (2009) Medico-legal findings, legal case progression, and outcomes in South African rape cases: retrospective review. PLoS Med 6: e1000164. doi:10.1371/journal.pmed.1000164.
75. Observatory for the Protection of Human Rights Defenders (2007) Steadfast in protest: annual report 2006. Paris: International Federation for Human Rights.
76. Fowler J (2006) Defending human rights and rebuilding with dignity. Washington (District of Columbia): US Holocaust Memorial Museum.
77. International Committee of the Red Cross (2011) Health care in danger: making the case. Geneva: International Committee of the Red Cross.

Editors' Summary

Background. Conflict in the Darfur region of Sudan between Arabic- and non-Arabic-speaking tribes over the past decade has resulted in a severe humanitarian crisis. According to the United Nations, more than 2.7 million people have fled from their homes to camps for internally displaced persons (IDPs) or to refugee camps in neighboring Chad, and up to 300,000 people have died from war, hunger, and disease since the conflict started. The origins of this conflict go back many years, but in 2003, organized rebel forces began attacking government targets, accusing the Government of Sudan (GoS) of oppressing black Africans in favor of Arabs. In response, the GoS attacked the rebel forces, but some observers allege it also targeted non-Arabic-speaking civilians, in contravention of international laws of war. Observers have also accused the GoS of having links with the Janjaweed militias, nomadic Arabs who attack settled black farmers, although the GoS denies any such links. Indeed, reports of systematic, targeted assaults on non-Arabic-speaking civilians, of large-scale disruption of rural livelihoods, and of deliberate consignment to living conditions likely to cause death have prompted some observers to accuse the GoS of genocide (violent crimes committed against a national, ethnical, racial, or religious group with the intention of destroying that group) and the International Criminal Court to issue arrest warrants for the allegedly responsible authorities.

Why Was This Study Done? Most investigations of claims of violence against civilians in Darfur have relied on self-reported data gathered from people living in refugee camps outside Sudan. Because these data could be biased, in this cross-sectional study (a study that characterizes a population at a single point in time), the researchers investigate the nature and geographic scope of alleged abuses against civilians in Darfur and endeavor to substantiate these allegations by analyzing the medical records of patients attending the Amel Centre for Treatment and Rehabilitation of Victims of Torture in Nyala, South Darfur. Opened in 2004, this center provided free clinical and legal services to civilians affected by human rights violations. Its staff fled in 2009 because of increasingly dangerous working conditions; the medical records used in this study were sent out of Sudan before the staff fled.

What Did the Researchers Do and Find? Between September 28, 2004, and December 31, 2006, 325 patients were seen at the Amel Centre. According to their medical records, 292 patients from 12 different non-Arabic-speaking tribes alleged that they had been attacked by GoS or Janjaweed forces in rural areas across Darfur. Nearly all the patients reported that they had been attacked in the absence of active armed conflict between GoS/Janjaweed forces and rebel groups. Half of them claimed that they had been beaten, two-fifths reported gunshot wounds, a third reported destruction or theft of property, and nearly a third reported involuntary detainment. Half of the 73 women seen at the center disclosed that they had been sexually assaulted, often near IDP camps. Only 198 medical records contained sufficient detail to enable the researchers to determine whether the documented medical evidence was consistent with the alleged abuses. However, in all these cases, the researchers judged that the medical evidence was consistent with, highly consistent with, or virtually diagnostic of the alleged abuses.

What Do These Findings Mean? These findings provide credible medical evidence that is consistent with torture and other human rights violations being inflicted on non-Arabic-speaking civilians in Darfur from 2004 to 2006. These findings cannot be used to estimate the population incidence of attacks on civilians or to corroborate claims of assailants' identities or of genocidal intent. Moreover, their accuracy may be affected by several limitations of this study. For example, during the study period, only patients who obtained a medical evidence form from the police were permitted to receive treatment from an authorized medical officer; obtaining such a form likely represented a considerable hurdle to accessing health care services. Nevertheless, the widespread, organized, and sustained pattern of attacks documented in this study is consistent with the possibility that the actions of Janjaweed and GoS forces during the conflict in Darfur may constitute war crimes, crimes against humanity, and/or acts of genocide. Importantly, these findings also highlight the need to provide adequate protection for health professionals working in countries affected by internal conflicts.

Additional Information. Please access these web sites via the online version of this summary at http://dx.doi.org/10.1371/journal.pmed.1001198.

- The African Union–United Nations Mission in Darfur (UNAMID) provides background information and up-to-date news about the ongoing conflict in Darfur Amnesty International, which campaigns for human rights, provides background information and news about the current situation in Darfur
- The Save Dafur Coalition also provides detailed information about the situation in Darfur Physicians for Human Rights, a non-profit organization that mobilizes health professionals to advance health, dignity, and justice, is calling for security in Darfur and compensation and restitution for survivors of the conflict
- Wikipedia has pages on Darfur and on genocide (note that Wikipedia is a free online encyclopedia that anyone can edit; available in several languages)
- Details on warrants of arrest issued by the International Criminal Court in response to the situation in Dafur are available

AIDS Care
2012, 1–8, iFirst

Hepatitis C virus infection is associated with painful symptoms in HIV-infected adults

Judith I. Tsui[a]*, Debbie M. Cheng[a,b], Howard Libman[c], Carly Bridden[a] and Jeffrey Samet[a,d]

[a]*Department of Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA, USA;* [b]*Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA;* [c]*Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA;* [d]*Department of Community Health Sciences, Boston University School of Public Health, Boston, MA, USA*

(*Received 22 March 2011; final version received 16 November 2011*)

The study aim was to assess whether hepatitis C virus (HCV) was associated with painful symptoms among patients with HIV. Using data from a prospective cohort of HIV-infected adults with alcohol problems, we assessed the effects of HCV on pain that interfered with daily living and painful symptoms (muscle/joint pain, headache and peripheral neuropathy). Exploratory analyses assessed whether depressive symptoms and inflammatory cytokines mediated the relationship between HCV and pain. HCV-infected participants ($n = 200$) had higher odds of pain that interfered with daily living over time (adjusted odds ratio [AOR] 1.43; 95% CI: 1.02–2.01; $p = 0.04$) compared to those not infected with HCV. HIV/HCV co-infected participants had higher odds of muscle or joint pain (AOR 1.45; 95% CI: 1.06–1.97; $p = 0.02$) and headache (AOR 1.57; 95% CI: 1.18–2.07; $p < 0.01$). The association between HCV and peripheral neuropathy did not reach statistical significance (AOR 1.33; 95% CI: 0.96–1.85; $p = 0.09$). Depressive symptoms and inflammatory cytokines did not appear to mediate the relationship between HCV and pain. Adults with HIV who are also co-infected with HCV are more likely to experience pain that interfered with daily living, muscle or joint pain, and headaches compared to those not co-infected. Research is needed to explore the association between HCV infection and pain, and to determine whether HCV treatment is an effective intervention.

Keywords: pain; hepatitis C; cytokines; symptoms; peripheral neuropathy

Introduction

Pain is a common condition in HIV-infected patients. A nationally representative survey conducted in 1996 reported that 67% of HIV-infected persons in the United States had experienced pain in the previous 4 weeks (Dobalian, Tsao, & Duncan, 2004). Pain is among the most common of symptoms reported by HIV-infected patients, and is strongly correlated with quality of life and psychological distress (Breitbart et al., 1998; Vogl et al., 1999). Substance abuse and unhealthy alcohol use, which commonly occur in the context of HIV-infection (Samet, Phillips, Horton, Traphagen, & Freedberg, 2004), may represent maladaptive coping responses to pain (Brennan, Schutte, & Moos, 2005; Holahan, Moos, Holahan, Cronkite, & Randall, 2001). A better understanding of the etiology of pain in HIV-infected individuals is needed to effectively prevent and treat pain in this population.

A co-morbidity whose contribution to pain among HIV-infected persons has been relatively unexplored is infection with hepatitis C virus (HCV). Chronic HCV infection has been associated with painful diagnoses such as peripheral neuropathy, arthritis, and arthralgias (Cacoub et al., 2000), as well as generalized pain in quality of life studies (Spiegel et al., 2005). Individuals with HCV have been noted to have a high prevalence of chronic pain and treatment for pain (Silberbogen, Janke, & Hebenstreit, 2007; Whitehead et al., 2008), and a study of substance users found that those who were HCV-infected were more likely to have persistent pain compared to those who were uninfected (Caldeiro et al., 2008). The majority of injection drug users are infected with HCV (Garfein, Vlahov, Galai, Doherty, & Nelson, 1996; Lorvick, Kral, Seal, Gee, & Edlin, 2001; Murrill et al., 2002; Thomas et al., 1995), and 15–30% of HIV-infected individuals are co-infected with HCV because of the shared risk factor of injecting drugs (Sulkowski & Thomas, 2003). Injection drug use (IDU) has been associated with greater pain in some studies of HIV-infected individuals (Del Borgo et al., 2001; Dobalian et al., 2004; Martin, Pehrsson, Osterberg, Sonnerborg, & Hansson, 1999; Riley et al., 2003), and underlying HCV infection may be an unrecognized contributory factor.

A number of causal mechanisms can be hypothesized to explain an association between HCV and pain. Depression, which is a known risk factor for pain in patients with (Richardson et al., 2009; Riley et al., 2003; Tsao, Dobalian, & Naliboff, 2004) and without HIV(Bair, Robinson, Katon, & Kroenke,

*Corresponding author. Email: judith.tsui@bmc.org

ISSN 0954-0121 print/ISSN 1360-0451 online

http://dx.doi.org/10.1080/09540121.2011.642989
http://www.tandfonline.com

2003), is a common co-morbidity among patients with HCV (Dwight et al., 2000; el-Serag, Kunik, Richardson, & Rabeneck, 2002; Fontana et al., 2002; Golden, O'Dwyer, & Conroy, 2005; Lee, Jamal, Regenstein, & Perrillo, 1997). Among HIV-infected patients, co-infection with HCV has been shown to be independently associated with depressive symptoms (Libman et al., 2006). HCV may cause a cytokine-induced depression; a small study demonstrated correlations between plasma levels of interleukin-1β (IL-1β) and tumor necrosis factor-α (TNF-α) and depressive symptoms (Loftis, Huckans, Ruimy, Hinrichs, & Hauser, 2008). Researchers have also suggested that HCV may predispose to chronic pain conditions such as fibromyalgia through a cytokine-mediated pathway (Thompson & Barkhuizen, 2003).

The primary aim of this study was to evaluate whether HCV status was associated with pain that interfered with daily living and painful symptoms in a cohort of HIV-infected adults with alcohol problems. A secondary aim was to explore whether depressive symptoms and serum levels of the inflammatory cytokines TNF-α, IL-6, and IL-10 mediated the relationship between HCV infection and pain.

Materials and methods

Design

Data were obtained from a prospective, observational cohort study (HIV-Longitudinal Interrelationships of Viruses and Ethanol [HIV-LIVE]), in which assessments occurred at 6-month intervals over a maximum of 48 months.

Subjects

Recruitment occurred from a previous cohort study, an intake clinic for HIV-infected patients, HIV primary care and specialty clinics at two hospitals, homeless shelters, drug treatment programs, subject referrals, and flyers. Enrollment occurred between August 2001 and July 2003. Eligibility criteria were as follows: (1) documented HIV antibody test by ELISA and confirmed by Western blot; (2) two or more affirmative responses to the CAGE alcohol screening questionnaire, or physician investigator diagnosis of alcoholism; and (3) ability to speak English or Spanish. Exclusion criteria included: (1) scoring <21 on the 30-item Folstein Mini-Mental State Examination (i.e., cognitive impairment); and (2) inability to provide informed consent. The Institutional Review Boards of Boston Medical Center and Beth Israel Deaconess Medical Center approved this study. Informed consent was obtained from each patient and the protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

Outcomes

The primary outcome was any pain interference with daily life, which was measured using a single question from the SF-12 (Ware, Kosinski, & Keller, 1996): "During the past 4 weeks, how much did pain interfere with your normal work (including housework)?" Pain interference was defined as any response other than "not at all." Secondary outcomes were musculoskeletal pain, headache, and peripheral neuropathy. These symptoms were measured using the HIV Symptom Index (Justice et al., 2001). This index is a validated 20-question inventory of symptoms common in the setting of HIV infection which assesses their frequency and consequences in the prior month. The questions asked: "How much have you been bothered by each of the following symptoms: 1) muscle aches or joint pain; 2) headache; and 3) pain, numbness or tingling in the hands and feet (peripheral neuropathy)?" Headache, muscle/joint pain, and peripheral neuropathy were defined as a response other than "I do not have this symptom." The primary and secondary outcomes were assessed at baseline and each follow-up visit.

Independent variables

The main independent variable was chronic HCV infection, as defined as a positive HCV antibody result confirmed with detectable HCV RNA level on polymerase chain reaction (PCR) testing. Participants who were HCV antibody-positive and HCV RNA-negative were considered not to have chronic HCV infection (i.e., they had cleared their infection either spontaneously or through treatment). HCV status was assessed only at baseline. Covariates included in the analyses were age, sex, race (black vs. non-black), current marital status (married vs. not married), current smoking, past 6 months homelessness past 6 months any IDU, past month any heavy alcohol use, low CD4 count (<200 cells/μL), HIV viral load (detectable vs. undetectable), HIV medication use in the past 6 months, depressive symptoms, and baseline serum levels of inflammatory cytokines TNF-α, IL-6, and IL-10. Hazardous alcohol use was defined as: (1) drinking greater than 14 standard drinks per week, or greater than 4 drinks in a day, for men; or (2) drinking greater than 7 drinks per week or greater than 3 drinks in a day for women (Saitz, 2005). Depressive symptoms were measured using the Center for Epidemiologic Studies Depression Scale (CES-D), and a threshold of ≥ 23 was used to define

substantial depressive symptoms (Golub et al., 2004). The cytokines TNF-α, IL-6, and IL-10 were chosen based on prior research demonstrating that (1) serum levels are significantly different in the setting of HCV infection compared to controls (Jacobson Brown & Neuman, 2001); and (2) they play a potential role in the pathogenesis of central pain conditions and chronic pain (Jacobson Brown & Neuman, 2001; Uceyler & Sommer, 2008; Verri et al., 2006; Watkins, Hutchinson, Milligan, & Maier, 2007). TNF-α and IL-6 were measured using Bio-Rad Luminex Flow Cytometry (Millipore) and IL-10 was measured using Chemiluminescent ELISA (R&D Systems). Laboratory testing was conducted at the University of Vermont's Laboratory for Clinical Biochemistry Research.

Statistical analyses

Chi-square and Student's *t*-tests were used to compare baseline characteristics of subjects with and without chronic HCV infection. General estimating equations (GEE) logistic regression was used to calculate odds ratios and 95% confidence intervals for each pain outcome. The GEE approach was used to account for the correlation from using repeated observations from the same subject over time. An exchangeable working correlation structure was used, and empirical standard errors are reported for all analyses. Collinearity of covariates was assessed by calculating the correlation between independent variables, and no pair of variables had a Spearman correlation >0.40. A two-tailed p-value <0.05 was considered statistically significant for all hypothesis testing. Final models were adjusted for all covariates. Covariates except for age, sex, and race were all modeled as time-dependent. Based on the approach described by Baron and Kinney (1986), exploratory analyses were performed to assess whether depressive symptoms and serum levels of cytokines TNF-α, IL-6, and IL-10 mediated the association between HCV infection and pain outcomes. As such, models that included and excluded depressive symptoms and inflammatory cytokines were compared to assess whether HCV coefficient estimates were attenuated by inclusion of the potential mediators. All subjects had data on depressive symptoms, however, only 309 had data on IL-6 and 343 had data on TNF-α and IL-10 due to insufficient serum sample for testing. Indicator variables were used to create a category for missing values for inflammatory cytokines, so that the same subset of respondents was examined in each analysis. All statistical analyses were conducted using SAS version 9.2 (SAS Institute, Inc., Cary, NC, USA).

Results

Of the 400 participants in the HIV-LIVE cohort, 397 had been tested for HCV RNA and were included in the analysis. Of the 397 HIV-infected subjects, 200 (50.4%) were found to have detectable HCV RNA and were considered to be co-infected. Participants who were HCV-infected were more likely to be older, homeless, have recently used injection drugs, diabetes, and have substantial depressive symptoms (Table 1). At baseline, the prevalence of having pain that interfered with daily living, muscle or joint pain, headache, and peripheral neuropathy was higher among HIV-infected participants who were co-infected with HCV compared to those who were not. HCV-infected participants were also significantly more likely to have baseline levels above the median for IL-6 and IL-10.

The median follow-up was 23.7 months (IQR: 12.2–31.7 months); the median number of follow-up assessments was 6 (IQR: 4 to 7). Analyses of longitudinal data included 2371 observations from 397 subjects. Results from final GEE logistic regression models (adjusted for time, age, sex, black race, marital status, homelessness, smoking, hazardous alcohol use, IDU, low CD4 count, detectable HIV viral load, use of HIV medications, and additional adjustment for diabetes for peripheral neuropathy outcome) are presented in row 1 of Table 2. Participants who were HCV-infected had significantly higher odds for experiencing pain that interfered with daily living over time (adjusted odds ratio [AOR] 1.43; 95% CI: 1.02–2.01) compared to those who were not HCV-infected. Likewise, HCV-infected participants were more likely to experience muscle or joint pain (AOR 1.45; 95% CI: 1.06–1.97) and headache (AOR 1.57; 95% CI: 1.18–2.07) over time compared to those who were not HCV-infected. The relative odds for HCV associated with peripheral neuropathy, while also greater than 1, did not reach statistical significance (AOR 1.33; 95% CI: 0.96–1.85).

The additional adjustment for substantial depressive symptoms somewhat attenuated the OR for HCV for all four pain outcomes (Table 2), suggesting that some of the association between HCV and pain may be explained by a greater burden of depressive symptoms in that group. However, the degree of attenuation was low ($<10\%$ in all models) suggesting depression is not a mediator of the relationship between HCV and pain. In the multivariate models, depressive symptoms were strongly associated with pain interference (AOR 1.99; 95% CI: 1.60–2.48; $p<0.01$), muscle or joint pain (AOR 2.35; 95% CI: 1.88–2.93; $p<0.01$), headache (AOR 2.25; 95% CI: 1.86–2.71; $p<0.01$), and

Table 1. Baseline characteristics of HIV-infected participants with and without HCV.

	HCV negative	HCV positive	
	($n = 197$)	($n = 200$)	p-value
Age (Mean/SD)	41 (± 8)	44 (± 7)	<0.01
Female	46 (23%)	54 (27%)	0.4
Black	87 (44%)	76 (38%)	0.21
Married	14 (7%)	13 (7%)	0.81
Smoker	145 (74%)	158 (79%)	0.21
Homeless[a]	40 (20%)	58 (29%)	0.04
Injection drug use[a]	9 (5%)	47 (24%)	<0.01
Heavy alcohol use[b]	66 (34%)	58 (29%)	0.32
Low CD4 count (<200 cells/μL)	33 (17%)	41 (21%)	0.4
Depressive symptoms (≥ 23 on CES-D)	72 (37%)	120 (60%)	<0.01
Diabetes	8 (4%)	20 (10%)	0.02
Pain that interferes with daily living	122 (62%)	144 (72%)	0.03
Muscle or joint pain	129 (66%)	153 (77%)	0.02
Headache	97 (49%)	132 (66%)	<0.01
Peripheral neuropathy	110 (56%)	136 (68%)	0.01
TNF-α (median/IQR)	6.5 (4.4–8.9)	7.6 (5.1–10.8)	
TNF-α above median	81 (46%)	90 (54%)	0.15
IL-6 (median/IQR)	2.1 (1.3–3.5)	3.4 (2.0–5.7)	
IL-6 above median	64 (40%)	90 (61%)	<0.01
IL-10 (median/IQR)	4.2 (2.5–6.6)	6.2 (4.0–9.3)	
IL-10 above median	68 (39%)	103 (62%)	<0.01

*Use in the past 6 months.
**Use in the past month.

peripheral neuropathy (AOR 1.54; 95% CI: 1.27–1.87; $p < 0.01$), and therefore, depressive symptoms may be a weak confounder. Levels of IL-6, IL-10, and TNF-α, however, were not significantly associated with any pain outcomes in multivariate models, and adjustment for these factors did not substantially impact the HCV effect in any model.

Discussion

In this cohort of HIV-infected persons with alcohol problems, co-infection with HCV was found to be associated with greater odds of experiencing pain that interfered with daily living, as well as symptoms of muscle or joint pain and headache. Although pain has been demonstrated to affect the majority of HIV-infected persons (Dobalian et al., 2004; Hewitt et al., 1997; Vogl et al., 1999), the mechanisms that underlie pain in this population are relatively poorly understood. Co-infection with HCV may be an overlooked and potentially modifiable risk factor for pain.

Substance use, and in particular IDU, has been associated with increased risk for pain at baseline and over time in HIV-infected patients (Del Borgo et al., 2001; Dobalian et al., 2004; Martin et al., 1999; Richardson et al., 2009; Riley et al., 2003; Tsao, Dobalian, & Stein, 2005). A study by Tsao et al. (2005) demonstrated that HIV-related disease burden mediated the association between illicit drug use and pain in a longitudinal cohort of HIV-infected persons. However, no published studies to date have examined the specific contributions of HCV co-infection to pain among persons living with HIV infection. Because HCV is strongly associated with IDU and has weaker associations with other drug use (Armstrong et al., 2006). it may be contributing to associations between substance use and pain observed in other studies. Our findings suggest that HCV may independently contribute to pain in HIV-infected patients. As prior studies have demonstrated the relative safety and efficacy of anti-HCV therapy among individuals with HIV (Carrat et al., 2004; Torriani et al., 2004), the potential to modify or prevent painful symptoms may provide additional motivation for addressing HCV in co-infected patients.

Our results demonstrating an association between HCV infection and muscle and joint pain are consistent with background studies in HIV-uninfected populations. Quality of life studies using the SF-36 have demonstrated greater bodily pain among patients who are HCV-infected versus HCV-uninfected (Spiegel et al., 2005). Clinical manifestations of HCV include extra-hepatic painful conditions. In a large

Table 2. Odds ratios for the association between HCV and pain outcomes.

	Pain that interferes with living			Muscle or joint pain			Headache			Peripheral neuropathy		
	OR	95% CI	*p*-value	OR	95% CI	*p*-value	OR	95% CI	*p*-value	OR	95% CI	*p*-value
Final model*	1.43	1.02–2.01	0.04	1.45	1.06–1.97	0.02	1.57	1.18–2.07	<0.01	1.33	0.96–1.85	0.09
Final model + depressive symptoms	1.34	0.96–1.88	0.09	1.33	0.98–1.81	0.07	1.44	1.09–1.90	0.01	1.27	0.92–1.77	0.15
Final model + depressive symptoms + high IL-6 (above median)	1.34	0.95–1.89	0.10	1.34	0.98–1.83	0.06	1.43	1.08–1.91	0.01	1.31	0.94–1.83	0.11
Final model + depressive symptoms + high IL-10 (above median)	1.29	0.91–1.83	0.15	1.27	0.94–1.74	0.12	1.42	1.07–1.88	0.01	1.29	0.93–1.79	0.12
Final model + depressive symptoms + high TNF-α (above median)	1.30	0.92–1.83	0.13	1.33	0.98–1.81	0.07	1.46	1.11–1.93	<.01	1.29	0.93–1.79	0.12

*GEE logistic regression models adjusted for time, age, sex, black race, marital status, homelessness, smoking, hazardous alcohol use, injection drug use, low CD4 count (<200), HIV viral load, and use of HIV medications; additional adjustment for diabetes for peripheral neuropathy outcome.

series of HCV-infected subjects, the prevalence of arthralgias and neuropathy was found to be 19% and 9%, respectively (Cacoub et al., 2000). A study of chronic liver disease patients found that HCV-infected patients had a higher relative odds for musculoskeletal pain compared to uninfected patients (Barkhuizen et al., 1999). To our knowledge, however, no prior studies have reported an association between HCV and headache. There is evidence, however, that HCV can be neuroinvasive (Laskus et al., 2005), and a large body of research demonstrates links between HCV and neurocognitive dysfunction (Perry, Hilsabeck, & Hassanein, 2008). The fact that HCV was associated with pain at multiple sites could suggest that infection with HCV affects one's overall tolerance of pain. Studies of pain tolerance using experimental techniques such as cold-pressor test have found differences in pain thresholds between opioid dependent former injection drug users and controls (Compton, Charuvastra, Kintaudi, & Ling, 2000; Compton, Charuvastra, & Ling, 2001; Doverty et al., 2001). It is conceivable that HCV status could, in part, be responsible for these differences.

We explored whether depressive symptoms and increased levels of inflammatory cytokines were mediators of the relationship between HCV and pain. Adjustment for depressive symptoms did not substantially attenuate the effects of HCV suggesting it is not a mediator. Results of the exploratory analyses also did not support that higher baseline levels of inflammatory cytokines were part of the causal pathway between HCV and pain. However, caution must be taken with interpreting these results as there were a number of limitations to the analyses. We measured serum levels of inflammatory cytokines at baseline only, and it is possible that changes in cytokine levels over time might better correlate with pain. Second, serum cytokine levels might not reflect the relevant physiologic pathway as cytokine expression can occur in other compartments (cerebrospinal fluid, peripheral blood mononucleocytes) (Loftis, Huckans, & Morasco, 2010).

There are additional limitations to this study. The main study outcome was based on a single question on the SF-12 questionnaire that reflected pain over the past 4 weeks. There was no information on pain severity, duration, or treatment available for analysis. Our findings could be related to undertreatment of pain among HCV-infected persons. Prior research has suggested that HIV-infected injection drug users are less likely to have pain effectively treated than those who do not use injection drugs (Breitbart et al., 1997). It is also possible that providers who care for patients with HCV infection may be reluctant to prescribe certain pain medications because of concerns about potential hepatotoxicity. The study was conducted on a population of HIV-infected adults with alcohol problem, which may affect the generalizability of its findings. However, given the prevalence of past alcohol problems among HIV-infected individuals, up to 40% in some studies (Samet et al., 2004), these findings would still be of importance even if not more broadly applicable. There may be other unmeasured confounders such as stress or anxiety associated with HCV infection that were not addressed in the analysis.

In summary, this study of HIV-infected adults with alcohol problems found that participants who were co-infected with HCV were more likely to experience pain that interfered with daily living,

muscle or joint pain, and headaches over time compared to those who were HCV-negative. Given that pain is a common morbidity among HIV-infected patients, more research is needed to explore the association between chronic HCV infection and pain in this population, and whether HCV treatment improves pain.

Acknowledgements

The authors thank Na Wang for her data management support. The authors report no institutional or commercial affiliations that might pose a conflict of interest. This study was supported by funding from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) R01-AA13216 and K24-AA015674 and the National Institute on Drug Abuse (NIDA) K23-DA027367.

References

Armstrong, G.L., Wasley, A., Simard, E.P., McQuillan, G.M., Kuhnert, W.L., & Alter, M.J. (2006). The prevalence of hepatitis C virus infection in the United States, 1999 through 2002. *Annals of Internal Medicine*, *144*(10), 705–714.

Bair, M.J., Robinson, R.L., Katon, W., & Kroenke, K. (2003). Depression and pain comorbidity: A literature review. *Archives of Internal Medicine*, *163*(20), 2433–2445.

Barkhuizen, A., Rosen, H.R., Wolf, S., Flora, K., Benner, K., & Bennett, R.M. (1999). Musculoskeletal pain and fatigue are associated with chronic hepatitis C: A report of 239 hepatology clinic patients. *American Journal of Gastroenterology*, *94*(5), 1355–1360.

Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Breitbart, W., Passik, S., McDonald, M.V., Rosenfeld, B., Smith, M., Kaim, M., & Funesti-Esch, J. (1998). Patient-related barriers to pain management in ambulatory AIDS patients. *Pain*, *76*(1–2), 9–16.

Breitbart, W., Rosenfeld, B., Passik, S., Kaim, M., Funesti-Esch, J., & Stein, K. (1997). A comparison of pain report and adequacy of analgesic therapy in ambulatory AIDS patients with and without a history of substance abuse. *Pain*, *72*(1–2), 235–243.

Brennan, P.L., Schutte, K.K., & Moos, R.H. (2005). Pain and use of alcohol to manage pain: Prevalence and 3-year outcomes among older problem and non-problem drinkers. *Addiction*, *100*(6), 777–786.

Cacoub, P., Renou, C., Rosenthal, E., Cohen, P., Loury, I., Loustaud-Ratti, V., … Piette, J.C. (2000). Extrahepatic manifestations associated with hepatitis C virus infection. A prospective multicenter study of 321 patients. The GERMIVIC. Groupe d'Etude et de Recherche en Medecine Interne et Maladies Infectieuses sur le Virus de l'Hepatite C. *Medicine (Baltimore)*, *79*(1), 47–56.

Caldeiro, R.M., Malte, C.A., Calsyn, D.A., Baer, J.S., Nichol, P., Kivlahan, D.R., & Saxon, A.J. (2008). The association of persistent pain with out-patient addiction treatment outcomes and service utilization. *Addiction*, *103*(12), 1996–2005.

Carrat, F., Bani-Sadr, F., Pol, S., Rosenthal, E., Lunel-Fabiani, F., Benzekri, A., … ANRS HCO2 RIBAVIC Study Team. (2004). Pegylated interferon alfa-2b vs standard interferon alfa-2b, plus ribavirin, for chronic hepatitis C in HIV-infected patients: A randomized controlled trial. *Journal of the American Medical Association*, *292*(23), 2839–2848.

Compton, P., Charuvastra, V.C., Kintaudi, K., & Ling, W. (2000). Pain responses in methadone-maintained opioid abusers. *Journal of Pain and Symptom Management*, *20*(4), 237–245.

Compton, P., Charuvastra, V.C., & Ling, W. (2001). Pain intolerance in opioid-maintained former opiate addicts: Effect of long-acting maintenance agent. *Drug and Alcohol Dependence*, *63*(2), 139–146.

Del Borgo, C., Izzi, I., Chiarotti, F., Del Forno, A., Moscati, A.M., Cornacchione, E., & Fantoni, M. (2001). Multidimensional aspects of pain in HIV-infected individuals. *AIDS Patient Care STDS*, *15*(2), 95–102.

Dobalian, A., Tsao, J.C., & Duncan, R.P. (2004). Pain and the use of outpatient services among persons with HIV: Results from a nationally representative survey. *Medical Care*, *42*(2), 129–138.

Doverty, M., White, J.M., Somogyi, A.A., Bochner, F., Ali, R., & Ling, W. (2001). Hyperalgesic responses in methadone maintenance patients. *Pain*, *90*(1–2), 91–96.

Dwight, M.M., Kowdley, K.V., Russo, J.E., Ciechanowski, P.S., Larson, A.M., & Katon, W.J. (2000). Depression, fatigue, and functional disability in patients with chronic hepatitis C. *Journal of Psychosomatic Research*, *49*(5), 311–317.

el-Serag, H.B., Kunik, M., Richardson, P., & Rabeneck, L. (2002). Psychiatric disorders among veterans with hepatitis C infection. *Gastroenterology*, *123*(2), 476–482.

Fontana, R.J., Hussain, K.B., Schwartz, S.M., Moyer, C.A., Su, G.L., & Lok, A.S. (2002). Emotional distress in chronic hepatitis C patients not receiving antiviral therapy. *Journal of Hepatology*, *36*(3), 401–407.

Garfein, R.S., Vlahov, D., Galai, N., Doherty, M.C., & Nelson, K.E. (1996). Viral infections in short-term injection drug users: The prevalence of the hepatitis C, hepatitis B, human immunodeficiency, and human T-lymphotropic viruses. *American Journal of Public Health*, *86*(5), 655–661.

Golden, J., O'Dwyer, A.M., & Conroy, R.M. (2005). Depression and anxiety in patients with hepatitis C: Prevalence, detection rates and risk factors. *General Hospital Psychiatry*, *27*(6), 431–438.

Golub, E.T., Latka, M., Hagan, H., Havens, J.R., Hudson, S.M., Kapadia, F., … Strathdee, S.A.

(2004). Screening for depressive symptoms among HCV-infected injection drug users: Examination of the utility of the CES-D and the Beck Depression Inventory. *Journal of Urban Health*, *81*(2), 278–290.

Hewitt, D.J., McDonald, M., Portenoy, R.K., Rosenfeld, B., Passik, S., & Breitbart, W. (1997). Pain syndromes and etiologies in ambulatory AIDS patients. *Pain*, *70*(2–3), 117–123.

Holahan, C.J., Moos, R.H., Holahan, C.K., Cronkite, R.C., & Randall, P.K. (2001). Drinking to cope, emotional distress and alcohol use and abuse: A ten-year model. *Journal of Studies on Alcohol*, *62*(2), 190–198.

Jacobson Brown, P.M., & Neuman, M.G. (2001). Immunopathogenesis of hepatitis C viral infection: Th1/Th2 responses and the role of cytokines. *Clinical Biochemistry*, *34*(3), 167–171.

Justice, A.C., Holmes, W., Gifford, A.L., Rabeneck, L., Zackin, R., Sinclair, G., ... Adult AIDS Clinical Trials Unit Outcomes Committee. (2001). Development and validation of a self-completed HIV symptom index. *Journal of Clinical Epidemiology*, *54*(Suppl. 1), S77–S90.

Laskus, T., Radkowski, M., Adair, D.M., Wilkinson, J., Scheck, A.C., & Rakela, J. (2005). Emerging evidence of hepatitis C virus neuroinvasion. *AIDS*, *19*(Suppl. 3), S140–S144.

Lee, D.H., Jamal, H., Regenstein, F.G., & Perrillo, R.P. (1997). Morbidity of chronic hepatitis C as seen in a tertiary care medical center. *Digestive Diseases and Sciences*, *42*(1), 186–191.

Libman, H., Saitz, R., Nunes, D., Cheng, D.M., Richardson, J.M., Vidaver, J., & Samet, J.H. (2006). Hepatitis C infection is associated with depressive symptoms in HIV-infected adults with alcohol problems. *American Journal of Gastroenterology*, *101*(8), 1804–1810.

Loftis, J.M., Huckans, M., & Morasco, B.J. (2010). Neuroimmune mechanisms of cytokine-induced depression: Current theories and novel treatment strategies. *Neurobiology of Disease*, *37*(3), 519–533.

Loftis, J.M., Huckans, M., Ruimy, S., Hinrichs, D.J., & Hauser, P. (2008). Depressive symptoms in patients with chronic hepatitis C are correlated with elevated plasma levels of interleukin-1beta and tumor necrosis factor-alpha. *Neuroscience Letters*, *430*(3), 264–268.

Lorvick, J., Kral, A.H., Seal, K., Gee, L., & Edlin, B.R. (2001). Prevalence and duration of hepatitis C among injection drug users in San Francisco, Calif. *American Journal of Public Health*, *91*(1), 46–47.

Martin, C., Pehrsson, P., Osterberg, A., Sonnerborg, A., & Hansson, P. (1999). Pain in ambulatory HIV-infected patients with and without intravenous drug use. *European Journal of Pain*, *3*(2), 157–164.

Murrill, C.S., Weeks, H., Castrucci, B.C., Weinstock, H.S., Bell, B.P., Spruill, C., & Gwinn, M. (2002). Age-specific seroprevalence of HIV, hepatitis B virus, and hepatitis C virus infection among injection drug users admitted to drug treatment in 6 US cities. *American Journal of Public Health*, *92*(3), 385–387.

Perry, W., Hilsabeck, R.C., & Hassanein, T.I. (2008). Cognitive dysfunction in chronic hepatitis C: A review. *Digestive Diseases and Sciences*, *53*(2), 307–321.

Richardson, J.L., Heikes, B., Karim, R., Weber, K., Anastos, K., & Young, M. (2009). Experience of pain among women with advanced HIV disease. *AIDS Patient Care STDS*, *23*(7), 503–511.

Riley, E.D., Wu, A.W., Perry, S., Clark, R.A., Moss, A.R., Crane, J., & Bangsberg, D.R. (2003). Depression and drug use impact health status among marginally housed HIV-infected individuals. *AIDS Patient Care STDS*, *17*(8), 401–406.

Saitz, R. (2005). Clinical practice. Unhealthy alcohol use. *New England Journal of Medicine*, *352*(6), 596–607.

Samet, J.H., Phillips, S.J., Horton, N.J., Traphagen, E.T., & Freedberg, K.A. (2004). Detecting alcohol problems in HIV-infected patients: Use of the CAGE questionnaire. *AIDS Research and Human Retroviruses*, *20*(2), 151–155.

Silberbogen, A.K., Janke, E.A., & Hebenstreit, C. (2007). A closer look at pain and hepatitis C: Preliminary data from a veteran population. *Journal of Rehabilitation Research and Development*, *44*(2), 231–244.

Spiegel, B.M., Younossi, Z.M., Hays, R.D., Revicki, D., Robbins, S., & Kanwal, F. (2005). Impact of hepatitis C on health related quality of life: A systematic review and quantitative assessment. *Hepatology*, *41*(4), 790–800.

Sulkowski, M.S., & Thomas, D.L. (2003). Hepatitis C in the HIV-infected person. *Annals of Internal Medicine*, *138*(3), 197–207.

Thomas, D.L., Vlahov, D., Solomon, L., Cohn, S., Taylor, E., Garfein, R., & Nelson, K.E. (1995). Correlates of hepatitis C virus infections among injection drug users. *Medicine (Baltimore)*, *74*(4), 212–220.

Thompson, M.E., & Barkhuizen, A. (2003). Fibromyalgia, hepatitis C infection, and the cytokine connection. *Current Pain and Headache Reports*, *7*(5), 342–347.

Torriani, F.J., Rodriguez-Torres, M., Rockstroh, J.K., Lissen, E., Gonzalez-Garcia, J., Lazzarin, A., ... APRICOT Study Group. (2004). Peginterferon Alfa-2a plus ribavirin for chronic hepatitis C virus infection in HIV-infected patients. *New England Journal of Medicine*, *351*(5), 438–450.

Tsao, J.C., Dobalian, A., & Naliboff, B.D. (2004). Panic disorder and pain in a national sample of persons living with HIV. *Pain*, *109*(1–2), 172–180.

Tsao, J.C., Dobalian, A., & Stein, J.A. (2005). Illness burden mediates the relationship between pain and illicit drug use in persons living with HIV. *Pain*, *119* (1–3), 124–132.

Uceyler, N., & Sommer, C. (2008). Cytokine regulation in animal models of neuropathic pain and in human diseases. *Neuroscience Letters*, *437*(3), 194–198.

Verri, W.A., Jr., Cunha, T.M., Parada, C.A., Poole, S., Cunha, F.Q., & Ferreira, S.H. (2006). Hypernociceptive role of cytokines and chemokines: Targets for analgesic drug development? *Pharmacology & Therapeutics*, *112*(1), 116–138.

Vogl, D., Rosenfeld, B., Breitbart, W., Thaler, H., Passik, S., McDonald, M., & Portenoy, R.K. (1999). Symptom prevalence, characteristics, and distress in AIDS outpatients. *Journal of Pain and Symptom Management*, *18*(4), 253–262.

Ware, J.E., Kosinski, M., & Keller, S.D. (1996). A 12-item short form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(4), 220–233.

Watkins, L.R., Hutchinson, M.R., Milligan, E.D., & Maier, S.F. (2007). "Listening" and "talking" to neurons: Implications of immune activation for pain control and increasing the efficacy of opioids. *Brain Research Reviews*, *56*(1), 148–169.

Whitehead, A.J., Dobscha, S.K., Morasco, B.J., Ruimy, S., Bussell, C., & Hauser, P. (2008). Pain, substance use disorders and opioid analgesic prescription patterns in veterans with hepatitis C. *Journal of Pain and Symptom Management*, *36*(1), 39–45.

JOURNAL OF PALLIATIVE MEDICINE
Volume 15, Number 7, 2012

DOI: 10.1089/jpm.2011.0505

A Randomized Controlled Trial of a Goals-of-Care Video for Elderly Patients Admitted to Skilled Nursing Facilities

Angelo E. Volandes, M.D.,[1,2] Gary H. Brandeis, M.D.,[3] Aretha Delight Davis, M.D., J.D.,[1,4] Michael K. Paasche-Orlow, M.D., M.A., M.P.H.,[3] Muriel R. Gillick, M.D.,[1,5] Yuchiao Chang, Ph.D.,[1,2] Elizabeth S. Walker-Corkery, M.P.H.,[2] Eileen Mann, R.N., M.S.N.,[1,2] and Susan L. Mitchell, M.D., M.P.H.[1,6]

Abstract

Objective: To determine the impact of a video on preferences for the primary goal of care.
Design, subjects, and intervention: Consecutive subjects 65 years of age or older ($n=101$) admitted to two skilled nursing facilities (SNFs) were randomized to a verbal narrative (control) or a video (intervention) describing goals-of-care options. Options included: life-prolonging (i.e., cardiopulmonary resuscitation), limited (i.e., hospitalization but no cardiopulmonary resuscitation), or comfort care (i.e., symptom relief).
Main measures: Primary outcome was patients' preferences for comfort versus other options. Concordance of preferences with documentation in the medical record was also examined.
Results: Fifty-one subjects were randomized to the verbal arm and 50 to the video arm. In the verbal arm, preferences were: comfort, $n=29$ (57%); limited, $n=4$ (8%); life-prolonging, $n=17$ (33%); and uncertain, $n=1$ (2%). In the video arm, preferences were: comfort, $n=40$ (80%); limited, $n=4$ (8%); and life-prolonging, $n=6$ (12%). Randomization to the video was associated with greater likelihood of opting for comfort (unadjusted rate ratio, 1.4; 95% confidence interval [CI], 1.1–1.9, $p=0.02$). Among subjects in the verbal arm who chose comfort, 29% had a do-not-resuscitate (DNR) order (κ statistic 0.18; 95% CI–0.02 to 0.37); 33% of subjects in the video arm choosing comfort had a DNR order (κ statistic 0.06; 95% CI–0.09 to 0.22).
Conclusion: Subjects admitted to SNFs who viewed a video were more likely than those exposed to a verbal narrative to opt for comfort. Concordance between a preference for comfort and a DNR order was low. These findings suggest a need to improve ascertainment of patients' preferences.
Trial Registration: Clinicaltrials.gov Identifier: NCT01233973.

Introduction

Elderly patients admitted to skilled nursing facilities (SNFs), are at a vulnerable point in their medical care.[1] These patients are recuperating from an acute illness and often have underlying complex chronic medical conditions. At the same time, SNF providers are likely to be unfamiliar with these patients' medical history and values. The hazards resulting from discontinuity of care between the hospital and SNF setting are well-described.[2–5]

Emerging health policy initiatives are increasingly focused on reducing avoidable, costly and burdensome rehospitalizations of older patients, particularly when patients' primary goal of care is comfort.[6,7] Optimizing advance care planning by ascertaining the patient's care preferences presents a key opportunity to promote goal-directed care that avoids unwanted and costly treatments.[8]

Over the last decade, an expanding body of evidence suggests that video decision aids help patients make better informed decisions by clarifying treatment options for a variety of life-limiting conditions including cancer and advanced dementia.[9–13] Video enhances communication beyond the usual *ad hoc* verbal approaches to advance care planning by providing realistic and standardized depictions of treatment options. To date, video support tools for advance care planning have not been studied in the post-acute care setting

[1]Harvard Medical School, Boston, Massachusetts.
[2]General Medicine Unit, Massachusetts General Hospital, Boston, Massachusetts.
[3]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts.
[4]Department of Medicine, Mt. Auburn Hospital, Cambridge, Massachusetts.
[5]Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, Massachusetts.
[6]Institute for Aging Research, Hebrew Senior Life, Boston, Massachusetts.
Accepted February 21, 2012.

where vulnerable older patients are often required to make value-laden decisions about treatment options.

To address this gap, we conducted a pilot randomized controlled trial of using a goals-of-care video for elderly patients admitted to a SNF after an acute hospitalization. We hypothesized that compared to subjects randomized to a verbal description of the goals of care, those viewing the video decision aid would be more likely to opt for treatments that align with comfort as the primary goal of care. An additional aim was to examine the correlation between a stated preference for comfort with the presence of a do-not-resuscitate (DNR) order in the medical record.

Methods

Subjects

The protocol was approved by the Partners Institutional Review Board and all subjects provided written informed consent. Subjects were recruited from a consecutive sample of patients admitted from acute hospitals to one of two SNFs in the greater Boston area. Both SNFs were located within nursing homes; one facility had 164 beds and the other had 190 beds.

Recruitment occurred between July 1, 2010 and February 28, 2011. At the time of admission, all patients who were 65 years of age or older and English-speaking were identified by the SNF admitting clerk and given a flier outlining the study. Patients were then asked by a SNF staff member (e.g., nurse or social worker) if they were interested in participating in the study. The final eligibility of patients who expressed a willingness to participate was then determined by a research assistant based on an assessment of their cognitive status performed within 72 hours of SNF admission. Patients were excluded if they had a Short Portable Mental Status Questionnaire (SPMSQ) score of less than 7.[14]

Design

All interviews were conducted in-person by one of two members of the research team (either A.E.V. or E.M.). Subjects underwent a baseline interview within 72 hours of admission to ascertain age, self-reported race, gender, religion, educational status, marital status, self-rated health status (excellent, very good, good, fair, or poor), and self-reported completion of advance directives (designated health care proxy or living will).

Immediately after the baseline interview, subjects were randomized into one of two decision-making modalities: (1) listening to a verbal narrative describing potential goals of care (control group) or (2) viewing a 6-minute video visually depicting treatments comprising the various potential goals of care (intervention group). We used simple randomization at each site based on a computer-generated scheme. Individual assignments were concealed in numbered envelopes. At the end of the trial, the randomization order of subjects was checked against the computer-generated list.

The video and verbal narratives were presented to the subjects in a quiet room in the SNF by a trained member of the research team who followed a structured script. Subjects randomized to the verbal control group were read a description of the three goals of medical care framework (life-prolonging care, limited medical care, and comfort care) developed in our previous work.[12,13,15,16] This framework was generated from a review of the advance care planning literature, and consultations with geriatric, critical care, palliative care, health literacy, and decision-making experts. Early versions of the framework were tested and validated with elderly subjects.[12,13,15,16]

The first option presented, life-prolonging care, was described as aiming to prolong life using all available medical care, and includes cardiopulmonary resuscitation and treatments in the intensive care unit. The second option presented, limited medical care, was described as aiming to maintain physical and mental functioning. It includes treatments such as hospitalization, intravenous fluids, antibiotics, but excludes cardiopulmonary resuscitation and treatments in the intensive care unit. The third option presented, comfort care, was described as aiming to maximize comfort and to relieve pain. It includes oxygen and analgesics, but excludes intravenous therapies and hospitalization unless necessary to provide comfort.

Subjects randomized to the intervention group viewed the video decision aid, shown on a portable computer. The 6-minute video depicts the three categories of medical care using the same definitions used in the verbal narrative, but includes visual images of the typical treatments comprising each goal. For example, life-prolonging care images includes: an intensive care unit with a ventilated patient being tended to by respiratory therapists; a simulated code on a mannequin with clinicians conducting cardiopulmonary resuscitation and intubation; and vasopressors administered through a venous catheter. Visual images to depict limited medical care include: a patient getting antibiotics via a peripheral intravenous catheter; scenes from a typical medical ward service; and a patient wearing a nasal cannula for oxygen delivery. The video depiction of comfort care includes: a patient on home hospice care receiving pain medications; a patient with a nasal cannula for oxygen at home; and, a medical attendant assisting a patient with self-care.

The development of the video decision aid followed a systematic approach, starting with a review of the advance care planning literature. The video's design, content, and structure were reviewed and edited for appropriateness and accuracy by geriatricians, critical care intensivists, palliative care physicians, and decision-making experts using an iterative process. The video was filmed without the use of prompts or stage directions to convey a candid realism in the style known as cinema verite.[17,18] All filming and editing was done by the investigative team (A.E.V. and A.D.D.) following previously published filming criteria.[19]

After exposure to either the verbal narrative or video, subjects were asked to select which level of care they would prefer if their medical condition worsened while at the SNF. Specifically, they were asked: "Imagine that you became very ill and in need of medical treatment, which general approach of medical care would you want provided: life-prolonging care, limited care, or comfort care?" Subjects unable to select a level of care were considered "uncertain."

For those subjects randomized to the video group, a four-point scale was used to assess perceived value of the video by asking subjects whether they were comfortable watching the video, whether they would recommend it to others, and whether they found the video helpful in their understanding of the goals-of-care options. Subjects who stated they did not

find the video helpful were asked to comment why, which was transcribed by the interviewer.

Last, a chart abstraction by the research assistant was performed immediately after exposure to the verbal or video narratives to ascertain whether or not the patient had a DNR order.

Statistical analysis

All subject characteristics and outcomes were described using proportions for categorical variables, and means and standard deviations (SD) for continuous variables. The primary outcome was a preference for comfort care (versus life prolongation, limited care or uncertain) as the goal of care and compared between the two groups using the exact χ^2 test. Additional analyses were conducted between subject characteristics (age, gender, education, marital status, health status, presence of advance directive, race, and randomization arm) and desire for comfort care for the entire cohort (i.e., both arms of the study). Rate ratios (RR) and 95% confidence interval (CI) were used to summarize the associations.

Finally, κ statistics were used to summarize the agreement between stated preferences for comfort and the presence of a DNR order in the medical record among the subjects in the intervention and control groups. A two-sided p value < 0.05 was considered statistically significant for all analyses.

With a target of 50 patients in each group, the power of the study was estimated to be 90% to detect a 30% difference in the preference for comfort care between the two groups. Data were analyzed and the randomization table was prepared using SAS software, version 9.2 (SAS Institute Inc., Cary, NC).

Results

Subject flow

A total of 155 consecutive English-speaking patients 65 years or older admitted to the two SNFs were approached to participate in the study, of whom 102 (66%) agreed to participate. Patients who declined did not differ significantly from the recruited subjects in terms of age, gender, or race. The most common reason given for not participating was lack of interest. Of the 102 patients expressing interest in the study, 1 patient was excluded because her SPMSQ score was less than 7, resulting in a total of 101 study subjects. A total of 51 subjects were randomized to the verbal control group, and 50 subjects were randomized to the video intervention group (Fig. 1). Baseline characteristics of the subjects are shown in Table 1.

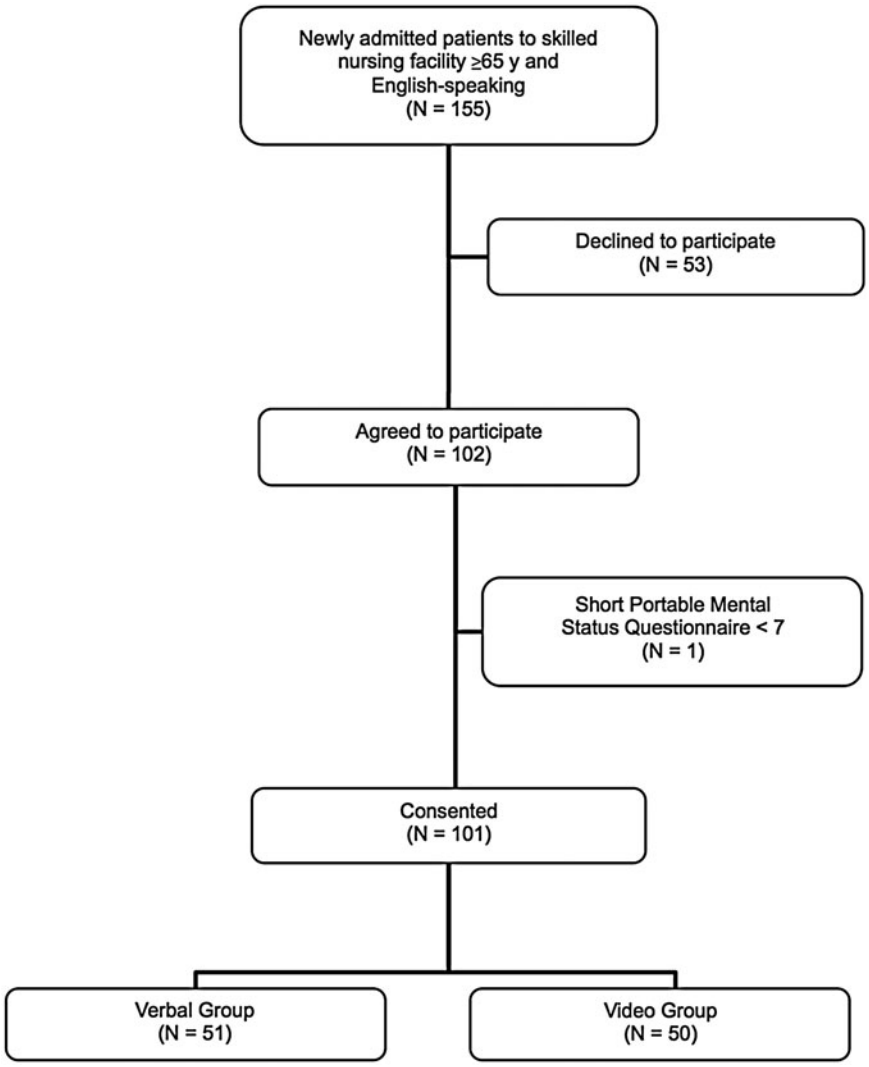


FIG. 1. Flow diagram of study and subjects' flow.

Outcomes

Among the 51 subjects receiving the verbal narrative, 29 (57%) preferred comfort care, 4 (8%) chose limited care, 17 (33%) desired life-prolonging care, and 1 (2%) was uncertain. Among the 50 subjects receiving the video decision aid, 40 (80%) chose comfort care, 4 (8%) chose limited care, and 6 (12%) desired life-prolonging care (Fig. 2). The proportion of subjects opting for comfort care was significantly higher in the intervention group ($p = 0.02$).

Subjects randomized to the video group had a greater likelihood of opting for comfort care (versus other responses; RR 1.4; 95% CI 1.1 to 1.9). None of the other subject characteristics were significantly associated with a preference for comfort care (Table 2).

In the verbal group, 28 of the 29 subjects who chose comfort care had a code status documented in their medical record. Only 8 of these 28 subjects (29%) had a DNR order in their medical record (κ statistic 0.18; 95% CI–0.02 to 0.37). In the video group, among the 40 subjects who chose comfort care, only 13 (33%) had a DNR order in their medical record (κ statistic 0.06; 95% CI–0.09 to 0.22). The video decision aid was highly acceptable to subjects in the intervention group: 45 of 50 (90%) subjects said they were "very comfortable" or "somewhat comfortable" viewing the video; 43 (86%) said they would "definitely" or "probably" recommend the video to others. Of the 50 subjects who viewed the video, only 4 (8%) found the video "not helpful." All 4 subjects who did not find the video helpful stated that they had already made their decisions previously and did not find viewing the video additionally helpful. There were no adverse events in either group.

Table 1. Characteristics of Older Subjects Randomized to the Verbal Narrative and Video Decision Aid Groups

Characteristics	*Verbal* (n = *51*)	*Video* (n = *50*)
Age, mean (SD), y	76 (7)	79 (9)
Women, *n* (%)	31 (61%)	30 (60%)
Race, *n* (%)		
Black or African American	33 (65%)	27 (54%)
Hispanic or Latino	2 (4%)	2 (4%)
White	16 (31%)	21 (42%)
Education, *n* (%)		
Elementary	8 (16%)	8 (16%)
Some high school	7 (14%)	7 (14%)
High school graduate	16 (31%)	16 (32%)
Some college	10 (20%)	7 (14%)
College graduate	6 (12%)	6 (12%)
Postgraduate or professional	4 (8%)	6 (12%)
Marital status, *n* (%)		
Married or with partner	17 (33%)	16 (32%)
Widowed	13 (25%)	17 (34%)
Divorced	6 (12%)	7 (14%)
Never married	15 (29%)	9 (18%)
Not answered	0	1 (2%)
Self-reported health status, *n* (%)		
Excellent	5 (10%)	5 (10%)
Very good	9 (18%)	5 (10%)
Good	14 (27%)	18 (36%)
Fair	11 (22%)	17 (34%)
Poor	12 (24%)	5 (10%)
Have an advance directive,[a] *n* (%)	28 (55%)	29 (58%)

[a]Subjects were asked if they had an advance directive, either a living will or health care proxy.
SD, standard deviation.

Discussion

Elderly subjects transitioning from an acute hospital stay to a SNF who view a video decision support tool for advance care planning are more likely to state comfort as a goal of care compared to those who listen to a verbal description about medical options. However, only about a third of subjects in both arms of the study who stated that they preferred comfort-oriented measures had a DNR order documented in their record. Thus, while a video decision aid increased the expressed desire for comfort care, our findings suggest a need to better translate those preferences into a written order to limit treatment.

As the first known randomized controlled trial of a video support tool to help determine the goals of care among elderly patients transitioning from an acute hospital to a SNF, this study supports and extends prior research about video decision-making tools in seriously ill patients.[12,13,15,16] In our previous work in advanced dementia and advanced cancer, video decision aids elicited preferences for comfort-oriented care, but these studies were conducted in outpatient settings and with patients who were making hypothetical decisions. Our current study extends this earlier work by showing the efficacy of the video for elderly patients in a clinical environment during periods of serious illness and transitioning across health care settings.

Delivery of medical care to patients that is consistent with their stated preferences is a critical consideration to providing high-quality medical care. The main finding in this study in a SNF is consistent with all earlier trials of video support tools for life-limiting conditions: subjects exposed to the video compared to a verbal narrative are more likely to opt for comfort-focused care (vs. life-prolonging or limited medical care). Our prior work also found that video decision support tools improve patient knowledge about their condition and treatment options, and reduce disparities among patients due poor health literacy.[12,13,20] Federal and state bodies will soon legislate the development of decision aids to assist patients and their families facing complex health care decisions.[21] Our work supports this initiative and may inform the type of decision support tools selected for further development and implementation.

Our findings show a lack of correlation between documented DNR status and stated preferences for comfort-oriented care regardless of decision-making modality. This suggests that in practice, more attention needs to be placed on providing patients and their families with opportunities to fully discuss preferences and to ensure these preferences are reflected in medical orders, especially during a vulnerable period as elderly patients transition between the hospital and post-acute care settings.[22–25] Ideally, an out-of-hospital DNR form would be completed that can travel with patients across sites.

Our study has several important limitations. First, the research staff collecting data were not blinded to randomization, which could introduce bias into our findings. Previous randomized studies of interventions aimed at improving end-of-life decision-making have seldom been blinded because limiting the number of interviewers eases the burden on

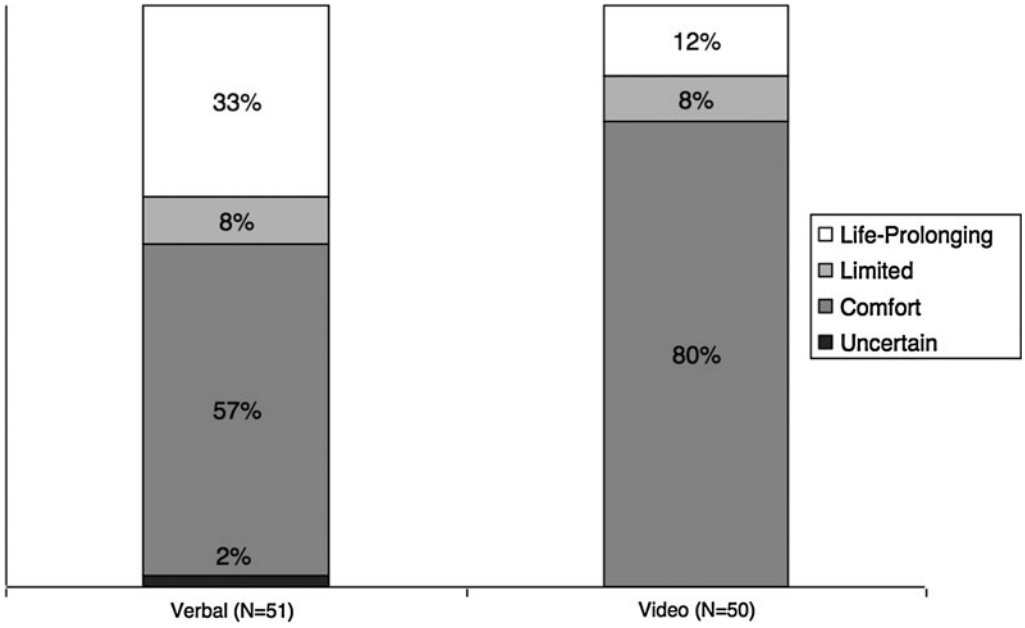


FIG. 2. Subjects' preferences for their goals of care.

subjects of addressing difficult subject matter.[26,27] We attempted to reduce the influence of this potential bias by using structured interviews and data collection instruments.

Second, one third of eligible subjects declined to participate in the study, although they did not differ significantly from the recruited subjects in terms of age, gender, or race. Third, videos can be manipulated to favor a particular perspective. Our study used one video exploring the goals of care. We did not assess responses to different videos altering the races and characteristics of the patients in the video. However, the video

Table 2. Analysis of the Likelihood of Choosing Comfort Care as the Primary Goal of Care as Opposed to Life-Prolonging or Limited Care

Characteristics	*Frequency in subjects choosing comfort care*	p *Value*	*Unadjusted rate ratio (95% CI)*
Age			
<80	40 (65%)		1.0 [reference]
≥80	29 (74%)	0.38	1.2 (0.9 to 1.5)
Gender, *n* (%)			
Female	45 (74%)		1.0 [reference]
Male	24 (60%)	0.19	0.8 (0.6 to 1.1)
Race, *n* (%)			
White	28 (76%)		1.0 [reference]
Black or [r:c4]African American	38 (63%)	0.26	0.8 (0.6 to 1.1)
Hispanic or Latino	3 (75%)	1.0	1.0 (0.5 to 1.8)
Education, *n* (%)			
Less than college graduate	52 (66%)		1.0 [reference]
College graduate or higher	17 (77%)	0.44	1.2 (0.9 to 1.5)
Marital status, *n* (%)			
Ever married	53 (70%)		1.0 [reference]
Never married	16 (67%)	0.80	1.0 (0.7 to 1.3)
Health status, *n* (%)			
Fair or poor	32 (71%)		1.0 [reference]
Good, very good, excellent	37 (66%)	0.67	0.9 (0.7 to 1.2)
Have advance directive, *n* (%)			
No	26 (60%)		1.0 [reference]
Yes	43 (75%)	0.13	1.2 (0.9 to 1.7)
Arm, *n* (%)			
Verbal	29 (57%)		1.0 [reference]
Video	40 (80%)	**0.02**	**1.4 (1.1 to 1.9)**

CI, confidence interval.

was evaluated by experts in a range of fields to ensure that it does not present a biased perspective or to otherwise try to influence subjects to choose a specific option. Fourth, we did not incorporate the video into clinical care by informing patients' physicians of their preferences and then following patients longitudinally to determine whether the video support tool had an effect on their advance care planning or actual clinical outcomes over time. This would be an ideal subsequent study to our present pilot study. Finally, our sample was limited to two SNFs in the Boston area. Thus, our findings might not be generalizable to elderly patients in other geographic areas or health care settings.

Discharge from an acute hospital stay to the post-acute care setting is a challenging and critical time of transition for older patients. Upon admission to a SNF, the patient's health status is often tenuous, and elucidation of their goals of care is needed to align ongoing treatments with patient preferences, and to help avoid unwanted and costly interventions for those whose goal of care is comfort. Our findings suggest that video decision aids may be a feasible and effective approach towards addressing this need in the setting of a SNF. However, our findings also demonstrate that ascertainment of goals of care may not be enough, as steps must also be taken to translate those wishes into a medical order (e.g., DNR orders).

In summary, elderly patients often face complex decision-making as they transition from the acute hospital setting to the SNF setting. To secure high-quality medical care at this juncture, patients must be informed about their options regarding end-of-life goals of care. Education of patients using video decision aids offers a more concrete portrait of potential goals of care compared to verbal discussions. As health care organizations look for innovative tools to inform patients at the end of life and to deliver high-quality medical care that is consistent with patient preferences, video decision aids may provide a useful tool. Future initiatives to improve ascertainment of patients' goals-of-care preferences should also include activities to improve the alignment between patient preferences, medical orders documenting patient goals-of-care, and the care that is delivered.

Acknowledgments

This research was supported by a grant from the Foundation for Informed Medical Decision Making. First author was supported by grant number K08HS018780 from the Agency for Healthcare Research and Quality. Dr. Mitchell was supported by grant K24AG033640 from the National Institute on Aging. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality or the National Institute on Aging. None of the funders participated in the collection, analysis, or interpretation of the data or in preparation, review, or approval of the manuscript.

Author Disclosure Statement

No competing financial interests exist.

References

1. Kane RL: Finding the right level of posthospital care: "We didn't realize there was any other option for him." JAMA 2011;305:284–293.
2. Coleman EA: Falling through the cracks: challenges and opportunities for improving transitional care for persons with continuous complex care needs. J Am Geriatr Soc 2003; 51:549–555.
3. Coleman EA, Berenson RA: Lost in transition: Challenges and opportunities for improving the quality of transitional care. Ann Intern Med 2004;141:533–536.
4. Coleman EA, Min SJ, Chomiak A, Kramer AM: Posthospital care transitions: patterns, complications, and risk identification. Health Serv Res 2004;39:1449–1465.
5. Coleman EA, Smith JD, Raha D, Min SJ: Posthospital medication discrepancies: Prevalence and contributing factors. Arch Intern Med 2005;165:1842–1847.
6. Mor V, Besdine RW: Policy options to improve discharge planning and reduce rehospitalization. JAMA 2011;305:302–303.
7. Mor V, Intrator O, Feng Z, Grabowski DC: The revolving door of rehospitalization from skilled nursing facilities. Health Aff 2010;29:57–64.
8. Berkowitz RE, Jones RN, Rieder R, et al: Improving disposition outcomes for patients in a geriatric skilled nursing facility. J Am Geriatr Soc 2011;59:1130–1136.
9. Barry MJ: Involving patients in medical decisions: How can physicians do better? JAMA 1999;282:2356–2357.
10. Entwistle V: The potential contribution of decision aids to screening programmes. Health Expect 2001;4:109–115.
11. Estabrooks C, Goel V, Thiel E, Pinfold P, Sawka C, Williams I: Decision aids: Are they worth it? A systematic review. J Health Serv Res Policy 2001;6:170–182.
12. El-Jawahri A, Podgurski LM, Eichler AF, et al: Use of video to facilitate end-of-life discussions with patients with cancer: A randomized controlled trial. J Clin Oncol 2010;28:305–110.
13. Volandes AE, Paasche-Orlow MK, Barry MJ, et al: Video decision support tool for advance care planning in dementia: Randomised controlled trial. BMJ 2009;338:b2159.
14. Pfeiffer E: A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc 1975;23:433–441.
15. Volandes AE, Barry MJ, Chang Y, Paasche-Orlow MK: Improving decision making at the end of life with video images. Medical Decision Making 2010;30:29–34.
16. Volandes AE, Ferguson LA, Davis AD, et al: Assessing end-of-life preferences for advanced dementia in rural patients using an educational video: A randomized controlled trial. J Palliat Med 2011;14:169–177.
17. Nichols B: *Introduction to Documentary*. Bloomington, IN: Indiana University Press; 2001.
18. Grant BK, Sloniowski J: *Documenting the Documentary: Close Readings of Documentary Film and Video.* Detroit, MI: Wayne State University Press; 1998.
19. Gillick MR, Volandes AE: The psychology of using and creating video decision aids for advance care planning. In: Lynch TE (ed): *Psychology of Decision Making in Medicine and Health Care*. New York, NY: Nova Science Publishers; 2007, pp. 193–206.
20. Volandes AE, Mitchell SL, Gillick MR, Chang Y, Paasche-Orlow MK: Using video images to improve the accuracy of surrogate decision-making: A randomized controlled trial. J Am Med Dir Assoc 2009;10:575–580.
21. The Patient Protection and Affordable Care Act, Public Law No: 111-148 (2010).
22. Coleman EA, Parry C, Chalmers S, Min SJ: The care transitions intervention: Results of a randomized controlled trial. Arch Intern Med 2006;166:1822–1828.

23. Coleman EA, Smith JD, Frank JC, Min SJ, Parry C, Kramer AM: Preparing patients and caregivers to participate in care delivered across settings: The Care Transitions Intervention. J Am Geriatr Soc 2004;52:1817–1825.
24. Naylor M, Brooten D, Jones R, Lavizzo-Mourey R, Mezey M, Pauly M: Comprehensive discharge planning for the hospitalized elderly. A randomized clinical trial. Ann Intern Med 1994;120:999–1006.
25. Naylor MD, Brooten D, Campbell R, et al: Comprehensive discharge planning and home follow-up of hospitalized elders: a randomized clinical trial. JAMA 1999;281:613–620.
26. SUPPORT: A controlled trial to improve care for seriously ill hospitalized patients. The study to understand prognoses and preferences for outcomes and risks of treatments (SUPPORT). The SUPPORT Principal Investigators. JAMA 1995;274:1591–1598.
27. Hofmann JC, Wenger NS, Davis RB, et al: Patient preferences for communication with physicians about end-of-life decisions. SUPPORT Investigators. Study to Understand Prognoses and Preference for Outcomes and Risks of Treatment. Ann Intern Med 1997;127:1–12.

Address correspondence to:
Angelo E. Volandes, M.D.
Massachusetts General Hospital
General Medicine Unit
50 Staniford Street
Boston, MA 02114

E-mail: avolandes@partners.org

Methadone Dose, Take Home Status, and Hospital Admission Among Methadone Maintenance Patients

Alexander Y. Walley, MD, MSc, Debbie M. Cheng, ScD, Courtney E. Pierce, MPH, Clara Chen, MHS, Tiffany Filippell, MD, Jeffrey H. Samet, MD, MA, MPH, and Daniel P. Alford, MD, MPH

Objectives: Among patients receiving methadone maintenance treatment (MMT) for opioid dependence, receipt of unobserved dosing privileges (take homes) and adequate doses (ie, ≥ 80 mg) are each associated with improved addiction treatment outcomes, but the association with acute care hospitalization is unknown. We studied whether take-home dosing and adequate doses (ie, ≥80 mg) were associated with decreased hospital admission among patients in an MMT.

Methods: We reviewed daily electronic medical records of patients enrolled in one MMT program to determine receipt of take-home doses, methadone dose 80 mg or more, and hospital admission date. Nonlinear mixed-effects logistic regression models were used to evaluate whether take-home doses or dose 80 mg or more on a given day were associated with hospital admission on the subsequent day. Covariates in adjusted models included age, sex, race/ethnicity, human immunodeficiency virus status, medical illness, mental illness, and polysubstance use at program admission.

Results: Subjects (n = 138) had the following characteristics: mean age 43 years; 52% female; 17% human immunodeficiency virus–infected; 32% medical illness; 40% mental illness; and 52% polysubstance use. During a mean follow-up of 20 months, 42 patients (30%) accounted for 80 hospitalizations. Receipt of take homes was associated with significantly lower odds of a hospital admission (adjusted odds ratio [AOR] = 0.26; 95% confidence interval [CI], 0.11-0.62), whereas methadone dose 80 mg or more was not (AOR = 1.01; 95% CI, 0.56-1.83).

Conclusions: Among MMT patients, receipt of take homes, but not dose of methadone, was associated with decreased hospital admission. Take-home status may reflect not only patients' improved addiction outcomes but also reduced health care utilization.

Key Words: dose, hospital admission, methadone maintenance treatment, take-home status

(*J Addict Med* 2012;6: 186–190)

Methadone maintenance treatment (MMT) in the United States is provided in federally regulated clinics that typically operate separate and in-parallel to mainstream medical care. Although methadone dosing should be individualized, as methadone metabolism exhibits individual variability (Lugo et al., 2005), higher doses are more effective than lower doses at retaining patients in treatment (Peles et al., 2010), diminishing craving, providing opioid blockade, and thus minimizing illicit opioid use and decreasing injection frequency (National Consensus Development Panel on Effective Medical Treatment of Opiate, Addiction, 1998; Strain et al., 1999; Faggiano et al., 2003; Center for Substance Abuse Treatment, 2005). National surveys of MMT programs have shown that the average methadone dose has been increasing since the 1980s and programs with higher average doses are more likely to be accredited by the Joint Commission (D'Aunno and Pollack, 2002; Pollack and D'Aunno, 2008). Thus, sufficiently high methadone dose may be one marker of MMT quality.

A "take home" is a dose of methadone given to the patient to take unobserved at home in place of requiring a return to the clinic a subsequent day for observed dosing. Contingent take-home doses are offered as rewards to patients with regular clinic attendance, counseling attendance, and abstinence from illicit drug use as measured typically by urine toxicology tests (Center for Substance Abuse Treatment, 2005). With adherence to MMT program expectations, take-home doses may increase up to a maximum of 6 or 13 consecutives doses, so eligible patients present to clinic every 1 or 2 weeks. Thus, contingent take-home dosing results from and reenforces success in MMT and is another potential marker of MMT quality. In a cohort study, it has been associated with longer retention in treatment and survival (Peles et al., 2011). Controlled studies have demonstrated that contingent take-home dosing increases abstinence from heroin and cocaine use (Chutuape et al., 1999, 2001) and increases counseling attendance (Kidorf et al., 1994).

From the Clinical Addiction Research and Education (CARE) Unit (AYW, DMC, JHS, DPA), Section of General Internal Medicine, Boston Medical Center, Boston, MA; Boston University School of Medicine (AYW, DMC, JHS, DPA), Boston, MA; Department of Biostatistics (DMC), Boston University School of Public Health, Boston, Massachusetts; Owen Graduate School of Management (CEP), Vanderbilt University, Nashville, TN; Data Coordinating Center (CC), Boston University School of Public Health, Boston, MA; and Department of Pediatrics (TF), Brown University School of Medicine, Providence, RI.

Received for publication August 14, 2011; accepted March 31, 2012.
Funded in part by the National Institute on Drug Abuse R25-DA13582.
The authors declare no conflicts of interest.

Send correspondence and reprint requests to Alexander Y. Walley, MD, MSc, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118. E-mail: awalley@bu.edu.

ISSN: 1932-0620/12/0603-0186
DOI: 10.1097/ADM.0b013e3182584772

As easily measurable markers of treatment adequacy and treatment success, methadone dose and take-home status may also reflect improved health care utilization. Patients receiving adequate methadone dosing and who have treatment success may be less likely to use acute hospital services, as they may be more capable of self-management (ie, attending to their medical needs). Addiction treatment generally is associated with decreased utilization of acute hospital services and, in some cases, lower overall costs (Laine et al., 2001, 2005; Weisner et al., 2001; Parthasarathy et al., 2003; Friedmann et al., 2006; Gourevitch et al., 2007). However, it is not clear which aspects of MMT are associated with improved medical care utilization.

In the setting of national health reform and the prospect of parity for the treatment of substance use disorders, the impact of MMT on the broader health system should be of interest to policymakers, providers, and patients. The determination of factors that reflect effective MMT will assist with evaluating health system outcomes. In this study, we reviewed the MMT and hospital electronic medical records (EMRs) of patients to determine longitudinally whether an association exists with hospital admission and 2 easily accessible clinical factors: (1) methadone dose and (2) take-home status.

METHODS

Study Design and Population

We conducted a retrospective medical record review of patients in the Boston Public Health Commission (BPHC) MMT program from February 1, 2006, to March 30, 2008. Only those individuals who met the following criteria were eligible for analysis: (1) enrolled in MMT on March 30, 2008; (2) received primary medical care from Boston Medical Center (BMC), an affiliated but physically and institutionally distinct medical center; and (3) had provided an active 2-way release of information with a BMC primary care physician. The release permitted ongoing exchange of medical information between the BPHC MMT program and providers at BMC for the purpose of clinical care coordination. This release was typically initiated at enrollment in MMT. For the purposes of the study, we linked patient data between the MMT program and the hospital. Investigators did not have direct contact with any study subjects as this study was a retrospective review of medical record data and therefore did not obtain any informed consents. This study was reviewed and approved by the institutional review board of Boston University and BMC, on behalf of the BPHC.

Data Collection and Measures

Outcomes

The primary outcome of the study was hospital admission (yes vs no) on a given day. To document hospital admissions, the medical record reviewers recorded start and end dates for each hospitalization listed in the MMT EMR from February 1, 2006, to March 30, 2008. The MMT nursing staff was informed by either hospital staff or the patient upon return about any hospitalization that resulted in a missed methadone dose. The nursing staff documented the dates, diagnoses, and locations of these hospitalizations in the MMT EMR after direct communication with the hospital staff. Forty-four of 55 hospitalizations (80%) reported to have occurred at BMC were confirmed in the BMC EMR. Periods when patients had no opportunity to receive take homes were excluded from analyses. These periods included the first 90 days of program admission, days incarcerated, and days hospitalized. Periods of known pregnancy were recorded and excluded because of expected hospitalizations during this time (pregnancy complications or child birth).

Independent Variables

There were 2 main independent variables in the study: methadone dose ($\geq$80 mg vs $<$80 mg) and receipt of take home (yes vs no) on a given day. Methadone dose and take-home status were extracted from the MMT EMR for each day the subject was enrolled from February 1, 2006, to March 30, 2008. Take homes are the primary incentive MMT programs use for positive reinforcement for patients who are succeeding in treatment and are granted after at least 90 days of complete methadone dosing attendance, counseling attendance, and no evidence of any illicit substances as determined by urine drug testing. Daily methadone doses were categorized as 80 mg or greater or less than 80 mg. This categorization was based on previous studies that demonstrated that doses 80 mg or more were associated with improved treatment outcomes (Strain et al., 1999). Both methadone dose and take-home status were modeled as time-dependent variables.

Covariates

From the BPHC EMR, reviewers recorded age, sex, race/ethnicity, and polysubstance use at the time of MMT admission. Race/ethnicity categories included 3 mutually exclusive categories—non-Hispanic white, non-Hispanic black/African American, and Hispanic. Subjects were categorized with polysubstance use if alcohol, cocaine, or benzodiazepine (nonprescribed) use was noted at methadone program admission. From the BMC EMR, we also recorded the active co-occurring medical and mental health conditions. Subjects were categorized as having medical illness if they had any of the following diagnoses documented: chronic obstructive pulmonary disorder, diabetes, renal disease, hypertension, cancer, pancreatitis, or hypercholesterolemia. Human immunodeficiency virus (HIV) status was documented and included as a covariate separately from other medical conditions. Subjects were categorized as having mental illness if they had a mood, thought, or anxiety disorder documented.

Analysis

Descriptive statistics of all subject characteristics at study entry were obtained and stratified by take-home status (ever vs never) and dose status (always $\leq$ 80 mg vs ever $\geq$ 80 mg). We used nonlinear mixed-effect logistic regression models to evaluate whether receipt of take homes or high-dose methadone were associated with a hospital admission on the following day. The mixed-effects regression model was used to incorporate multiple observations available from the same individual (eg, subjects could have repeated hospitalizations) in the analysis while controlling for the correlation within individuals to obtain proper estimates of variability. The

primary analysis included both independent variables in the same adjusted model. Because some previous studies define 60 mg of methadone as "high dose" (Faggiano et al., 2003), we repeated the model, using a definition of 60 mg or higher in a sensitivity analysis. Covariates in the adjusted model included age, sex, race/ethnicity, HIV status, chronic medical illness, mental illness, polysubstance use, length of time in treatment before study enrollment, and time since baseline. To minimize the potential for collinearity, we assessed correlation between pairs of independent variables and verified that no pair of variables included in the same regression model was highly correlated (ie, $r > 0.40$). Analyses were conducted using 2-sided tests and a significance level of 0.05. All analyses were performed using SAS software version 9.1 (SAS Institute Inc., Cary, North Carolina).

RESULTS

Of the 365 patients enrolled in the MMT program, 71 (19%) had no release signed for communication with a primary care physician and 128 (35%) had a signed release, but not with a physician at the affiliated medical center. Of the remaining 166 subjects screened, 28 were excluded because they had been enrolled in MMT less than 90 days as of March 30, 2008. Hence, 138 of the 365 patients (39%) were ultimately found eligible and included in study analyses (Figure 1).

The following were patient characteristics at study entry: mean age of 43 years; 52% female; 49% white, 29% black or African American, and 22% Hispanic; 17% HIV-infected; 32% chronic medical illness; 40% mental illness; and 52% polysubstance use. Table 1 details patient characteristics overall and stratified by both take-home status and methadone dose ($\geq$80 mg or $<$80 mg). Among the study patients, 52% had at least 1 documented take home during the follow-up period and 75% had at least 1 documented methadone dose of 80 mg or more. Those patients who ever received take homes were less likely to have polysubstance use at baseline, were older, and had more follow-up months during the study and more months in treatment before the study. Those patients with a dose always 80 mg or less were less likely to have mental illness or polysubstance use and had fewer follow-up months during the study. Among the 138 subjects, the total number of eligible observed days in MMT was 83,149 and the mean duration of follow-up time was 20 months. Of the 80 hospital admissions among 42 (30%) subjects, 9 (11%) admissions were among subjects with take homes and 50 (62.5%) admissions were among subjects with 80 mg or more of methadone. The overall hospitalization rate was 0.35 hospitalizations per person-year of eligible treatment days.

In models adjusted for age, sex, race/ethnicity, mental and medical illness, HIV infection, and polysubstance use, receipt of take homes was associated with significantly lower odds of a hospital admission (odds ratio = 0.26; 95% confidence interval [CI], 0.11-0.62), whereas dose of 80 mg or more was not (odds ratio = 1.01; 95% CI, 0.56-1.83) (Table 2). In a sensitivity analysis, methadone dose of 60 mg or more was also not significantly associated with hospital admission. In both models, medical illness was associated with increased odds of hospital admission. Common hospitalization diagnoses were typical for that seen in an urban hospital and are listed in Table 3.

DISCUSSION

Among MMT patients with established medical care, those who demonstrated treatment success in MMT as defined by achieving unobserved dosing privileges (ie, take home doses) had one quarter the odds of hospital admission

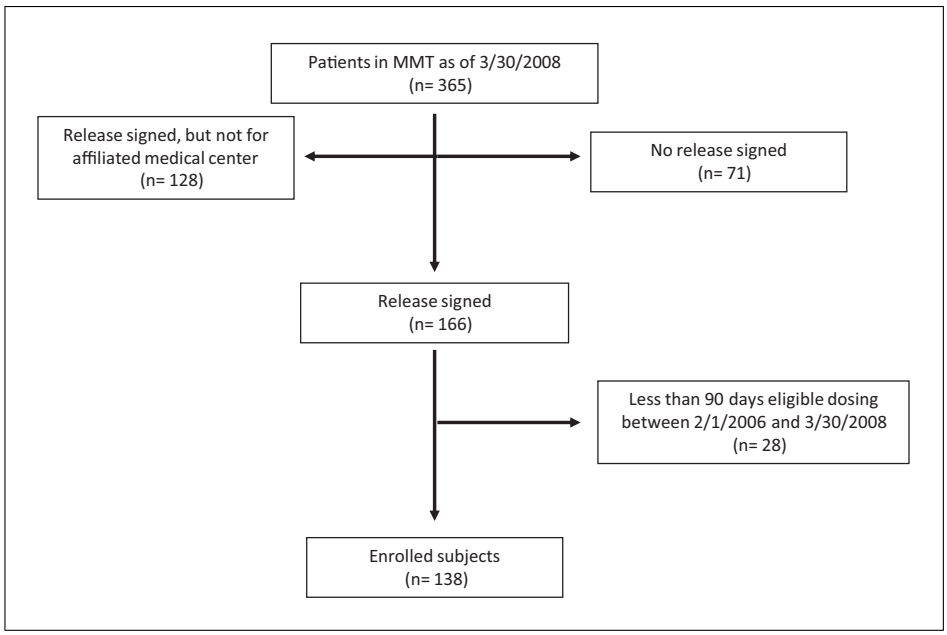


FIGURE 1. Identification of the study sample among methadone maintenance treatment patients. MMT indicates methadone maintenance treatment.

TABLE 1. Subject Characteristics at Study Enrollment (n = 138)

	Total, n = 138	Take Homes ever, n = 72 (52%)	Never Take Homes, n = 66 (48%)	*P*	Dose < 80 mg Always, n = 35 (25%)	Dose ≥ 80 mg Ever, n = 103 (75%)	*P*
Male, n (%)	66 (48)	39 (54)	27 (41)	0.1	14 (40)	52 (50)	0.3
Race/ethnicity, n (%)				0.3			0.6
Non-Hispanic white	68 (49)	34 (47)	34 (52)		17 (49)	51 (50)	
Non-Hispanic black	40 (29)	19 (26)	21 (32)		12 (34)	28 (27)	
Hispanic	30 (22)	19 (26)	11(17)		6 (17)	24 (23)	
HIV infection, n (%)	23 (17)	13 (18)	10 (15)	0.6	4 (11)	19 (18)	0.3
Medical illness, n (%)	44 (32)	26 (36)	18 (27)	0.3	10 (29)	34 (33)	0.6
Mental illness, n (%)	55 (40)	27 (38)	28 (42)	0.6	8 (23)	47 (46)	0.02
Polysubstance use, n (%)	72 (52)	31 (43)	14 (62)	<0.01	12 (34)	60 (58)	0.01
Age, mean (SD)	43 (11.5)	46 (11.4)	39 (10.5)	<0.01	42 (12.4)	43 (11.2)	0.8
Mean dose, mean (SD)	85 (30.9)	90 (33.1)	80 (27.7)	0.06	50 (12.9)	97 (26.0)	<0.01
Follow-up months, mean (SD)	20 (8.0)	23 (5.2)	16 (9.0)	<0.01	17 (9.6)	21 (7.1)	0.02
Months in treatment before study, mean (SD)	39 (54.7)	60 (64.9)	17 (27.4)	<0.01	47 (67.3)	37 (49.9)	0.4

HIV, human immunodeficiency virus.

TABLE 2. Multivariate Nonlinear Mixed-effects Logistic Regression Model Evaluating Association Between Take-home Status, Methadone Dose, and Hospital Admission (n = 138)*

	Adjusted Odds Ratio	95% Confidence Interval	*P*
Take-home status	**0.26**	**0.11-0.62**	**<0.01**
Dose ≥ 80 mg	1.01	0.56-1.83	0.97
Age (per 1 yr)	1.01	0.97-1.05	0.70
Gender			
Female	Ref		
Male	1.25	0.64-2.44	0.51
Race/ethnicity			
White	Ref		
Black/African American	0.67	0.29-1.52	0.33
Hispanic	0.73	0.31-1.72	0.47
HIV	1.77	0.82-3.81	0.14
Medical illness	2.41	1.20-4.85	0.01
Mental illness	0.61	0.31-1.22	0.16
Polysubstance use	1.28	0.63-2.61	0.49

*Model also adjusted for follow-up time as a time-varying covariate and months in treatment before study.
HIV, human immunodeficiency virus.

TABLE 3. Primary Diagnoses for Hospitalizations of Methadone Maintenance Patients

	N = 80	%
Pneumonia or upper respiratory infection	11	13.8
Cardiac (chest pain, arrhythmia)	9	11.2
Gastrointestinal (pancreatitis, diverticulitis, melena, abdominal pain, biliary obstruction)	8	10
Other infection (viral, bacteremia, urinary tract, osteomyelitis)	7	8.8
Asthma or chronic obstructive pulmonary disease exacerbation	6	7.5
Trauma or musculoskeletal pain	6	7.5
Psychiatric	5	6.2
Diabetes	4	5
Soft tissue infection (cellulitis)	3	3.8
Renal failure	2	2.5
Altered mental status (delirium, encephalopathy)	2	2.5
Paralysis agitans	1	1.2
Spontaneous ecchymosis	1	1.2
Detoxification	1	1.2
Release from incarceration	1	1.2
Unknown	13	16.2

compared with those not receiving take-home doses. Thus, achieving take-home doses represented not only substantial addiction treatment success but also a marked decreased risk of medical hospitalization during this period. Although this analysis does not determine whether take-home status directly reduces hospitalization or is a marker of other unmeasured factors, it does account for other known predictors of hospital admission, including age, HIV infection, chronic medical illnesses, and mental illness.

We found no evidence that the dose of methadone was associated with decreased hospitalization. The observed lack of association between doses greater than 80 mg and hospitalization suggests that the benefits of a higher dose do not include decreased hospital admissions.

The Office of National Drug Control Policy 2010 Strategy highlights increasing quality and performance within addiction treatment as a key objective and using quality and performance measures to improve addiction treatment (Office of National Drug Control Policy, 2010). The treatment of opioid dependence with pharmacotherapy (eg, MMT) is one of the 11 evidence-based practices identified by the National Quality Forum and cited by the Office of National Drug Control Policy for widespread adoption. Providing pharmacotherapy for opioid dependence should include demonstrating treatment quality and performance. Hospitalization is an important outcome to consider because of both health and cost implications. Adequate MMT dosing is already recognized as a measure of MMT effectiveness in that lower doses result in poorer addiction treatment outcomes (National Consensus Development Panel on Effective Medical Treatment of Opiate, Addiction, 1998; Strain et al., 1999; Faggiano et al., 2003; Center for Substance Abuse Treatment, 2005). Take-home status has been less studied, but by definition it is a marker of treatment success, because abstinence from drugs, participation in counseling, and daily dosing are required to receive take homes. However, whether or not such success extends to utilization of medical resources has not been examined in previous studies. By demonstrating an association between take-home status

and hospitalization, an important medical and health system cost outcome, this study supports consideration of take-home status as a useful performance measure reflecting direct benefit to individual patients and the wider health care system.

This study's strengths included the use of the daily MMT program EMR, which allowed us to determine each patient's methadone dose, take-home status, and hospitalization status for every day they were enrolled in MMT during the study period. By including subjects with a documented primary care physician and a signed release of information, our sample allowed for documentation of medical problems through the EMR at the affiliated medical center. Although this inclusion criterion limited the proportion of eligible patients included in the analysis and limits the generalizability of our findings, it is unlikely that this would bias the results, because engagement in primary care likely decreases acute care utilization. This study was also limited by the nature of retrospective design and focus on a single clinic. However, because MMT is subject to federal regulations, there are substantial similarities from clinic to clinic.

CONCLUSIONS

Among MMT patients, receipt of take homes, but not dose of methadone, was significantly associated with reduced medical hospitalizations. Thus, take-home status reflects not only patients' improved addiction outcomes but also reduced health care utilization. Given that this characteristic is easy to measure, clinically important and had cost implications, it should be further considered as a useful quality and performance measure for MMTs. Furthermore, intervention studies that seek to increase take-home status among MMT patients should evaluate health care utilization outcomes.

ACKNOWLEDGMENTS

The authors thank the staff at the Opioid Treatment Program, Boston Public Health Commission, for their support and cooperation during this study. This project was funded in part by the National Institute on Drug Abuse R25-DA13582. Parts of this work were presented at the Society of General Internal Medicine annual meeting May 13-16, 2009, and the College on Problems in Drug Dependence annual meeting June 24, 2009.

REFERENCES

Center for Substance Abuse Treatment. Medication-Assisted Treatment for Opioid Addiction in Opioid Treatment Programs. Treatment Improvement Protocol (TIP) Series 43. 2005. DHHS Publication Number (SMA) 05-4048. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2005.

Chutuape MA, Silverman K, Stitzer ML. Effects of urine testing frequency on outcome in a methadone take-home contingency program. *Drug Alcohol Depend* 2001;62:69–76.

Chutuape MA, Silverman K, Stitzer ML. Use of methadone take-home contingencies with persistent opiate and cocaine abusers. *J Subst Abuse Treat* 1999;16:23–30.

D'Aunno T, Pollack HA. Changes in methadone treatment practices: results from a national panel study, 1988–2000. *JAMA* 2002;288:850–856.

Faggiano F, Vigna-Taglianti F, Versino E, Lemma P. Methadone maintenance at different dosages for opioid dependence. *Cochrane Database Syst Rev* 2003;(3):002208.

Friedmann PD, Hendrickson JC, Gerstein DR, Zhang Z, Stein MD. Do mechanisms that link addiction treatment patients to primary care influence subsequent utilization of emergency and hospital care? *Med Care* 2006;44: 8–15.

Gourevitch MN, Chatterji P, Deb N, Schoenbaum EE, Turner BJ. On-site medical care in methadone maintenance: associations with health care use and expenditures. *J Subst Abuse Treat* 2007;32:143–151.

Kidorf M, Stitzer ML, Brooner RK, Goldberg J. Contingent methadone take-home doses reinforce adjunct therapy attendance of methadone maintenance patients. *Drug Alcohol Depend* 1994;36:221–226.

Laine C, Hauck WW, Gourevitch MN, Rothman J, Cohen A, Turner BJ. Regular outpatient medical and drug abuse care and subsequent hospitalization of persons who use illicit drugs. *JAMA* 2001;285:2355–2362.

Laine C, Lin YT, Hauck WW, Turner BJ. Availability of medical care services in drug treatment clinics associated with lower repeated emergency department use. *Med Care* 2005;43:985–995.

Lugo RA, Satterfield KL, Kern SE. Pharmacokinetics of methadone. *J Pain Palliat Care Pharmacother* 2005;19:13–24.

National Consensus Development Panel on Effective Medical Treatment of Opiate, Addiction. Effective medical treatment of opiate addiction. *JAMA* 1998;280:1936–1943.

Office of National Drug Control Policy. National Drug Control Strategy: 2010. Washington, DC: Office of National Drug Control Policy; 2010. http://www.whitehouse.gov/sites/default/files/ondcp/policy-and-research/ndcs2010.pdf. Accessed May 23, 2012.

Parthasarathy S, Mertens J, Moore C, Weisner C. Utilization and cost impact of integrating substance abuse treatment and primary care. *Med Care* 2003;41:357–367.

Peles E, Schreiber S, Adelson M. 15-Year survival and retention of patients in a general hospital-affiliated methadone maintenance treatment (MMT) center in Israel. *Drug Alcohol Depend* 2010;107:141–148.

Peles E, Schreiber S, Sason A, Adelson M. Earning "take-home" privileges and long-term outcome in a methadone maintenance treatment program. *J Addict Med* 2011;5:92–98.

Pollack HA, D'Aunno T. Dosage patterns in methadone treatment: results from a national survey, 1988–2005. *Health Serv Res* 2008;43:2143–2163.

Strain EC, Bigelow GE, Liebson IA, Stitzer ML. Moderate- vs high-dose methadone in the treatment of opioid dependence: a randomized trial. *JAMA* 1999;281:1000–1005.

Weisner C, Mertens J, Parthasarathy S, Moore C, Lu Y. Integrating primary medical care with addiction treatment: a randomized controlled trial. *JAMA* 2001;286:1715–1723.

Journal of Substance Abuse Treatment 43 (2012) 377–381

Contents lists available at SciVerse ScienceDirect

Journal of Substance Abuse Treatment

Integration of substance use treatment and medical care: A special issue of *JSAT*

Alexander Y. Walley, M.D., M.Sc. [a], Jeanette M. Tetrault, M.D. [b], Peter D. Friedmann, M.D., M.P.H. [c,*]

[a] *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA, 02118, USA*
[b] *Department of Internal Medicine, Section of General Internal Medicine, Yale University School of Medicine, New Haven, CT,06510, USA*
[c] *Research Enhancement Award Program, Research Service, Department of Veteran Affairs Medical Center; Division of General Internal Medicine, Rhode Island Hospital; Departments of Medicine and Health Services, Policy & Practice, Alpert Medical School of Brown University, Providence, RI 02903, USA*

ARTICLE INFO

Article history:
Received 7 September 2012
Accepted 7 September 2012

1. Introduction

Integration of substance use treatment and medical care has long promised to enhance the care of patients with alcohol and drug abuse, and related medical conditions (Samet, Friedmann, & Saitz, 2001). In 2006, the Institute of Medicine called for normalization of collaboration and coordination among health care providers that serve patients with medical, mental health and substance use disorders (Committee on Crossing the Quality Chasm, 2006). The integration of these services could improve the quality and efficiency of care by diagnosing and treating more than one condition at once. Integration can be *good for patients* with more than one condition, because it allows for "one stop shopping" resulting in fewer visits to fewer places for care. The potential for longitudinal relationships to develop over time may result in increased understanding and trust, essential components in effective care for substance use disorders. Integration can also be *good for providers* because it facilitates team-based care where no individual provider takes on the full burden of a patient's multiple conditions, but the work is shared among multiple disciplines. Communication and collaboration can be a daily part of the team culture, eliminating an extra step to coordinate care.

In the fall of 2011, *JSAT* recruited the co-chairs of the Substance Abuse Interest Group of the Society of General Internal Medicine, Dr. Jeanette Tetrault and Dr. Alexander Walley, to serve as guest editors of a special issue to advance the science in this area. A call for submissions was issued in January 2012. Original research addressing the focus of the special issue received *JSAT*'s standard peer review: two experts and an editor assessed each paper's pertinence and scholarly merit. The papers included in this special issue address a range of topics and settings that are relevant to the integration of medical care and substance use treatment.

This diverse group of articles demonstrates that interventions for the detection and management of substance-related disorders can reach patients in medical settings. The converse is also true: the integration of medical care into specialty addiction treatment improves the effectiveness of care for substance-related medical conditions such as hepatitis C and HIV. Some important themes that emerge from this collection of papers enhance our understanding as to when and how these integrated models work. Several of the articles address the challenges of delivering high-quality interventions in medical settings. Others address the critical issue of complex comorbidity, as screening in medical settings will detect patients with substance use and co-occurring medical and psychiatric conditions for whom adequate treatment will require more than brief intervention. Fortunately, effective models exist for delivering care to patients with comorbid needs, of which some examples appear in this issue. Finally, this collection demonstrates the potential of pharmacotherapy as a tool for the innovative treatment of more severe substance use problems by medical providers.

2. Delivering high-quality care in integrated medical settings

2.1. Quality of brief interventions delivered in medical settings

Screening, Brief Intervention, and Referral to Treatment (SBIRT) programs have received nationwide attention and a growing body of literature describes successful ways to implement SBIRT in clinical settings (Substance Abuse & Mental Health Services Administration, 2012). Brief interventions (BI) for at-risk alcohol use have proven

* Corresponding author.
E-mail address: pfriedmann@lifespan.org (P.D. Friedmann).

0740-5472/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsat.2012.09.002

efficacy in a variety of clinical settings (Academic, 2007; Babor et al., 2007; Estee, Wickizer, He, Shah, & Mancuso, 2010; Fleming, Barry, Manwell, Johnson, & London, 1997) and BI delivered in the emergency department can reduce both drinking outcomes and health care-related cost (D'Onofrio, Pantalon, Degutis, et al., 2008; Gentilello, Ebel, Wickizer, Salkever, & Rivara, 2005; Havard, Shakeshaft, & Sanson-Fisher, 2008). BI is a "teachable skill" that can be readily integrated into existing graduate medical education curricula (Bernstein, Bernstein, Feldman, et al., 2007; D'Onofrio, Nadel, Degutis, et al., 2002; Ryan, Martel, Pantalon, et al., 2012; Tetrault, Green, Martino, et al., 2012); however, no formal assessment tool exists to measure the degree to which practitioners adhere to the key components of BI. Pantalon et al., 2012 describe the development and psychometric testing of a tool to measure practitioner adherence to a type of brief intervention called the Brief Negotiation Interview (BNI) (Pantalon). The BNI Adherence Scale was originally used in the context of a randomized clinical trial of BNI for hazardous drinking in the emergency department. Through review of 342 audio-taped patient encounters in the emergency department, the authors found that the BNI Adherence Scale had excellent internal consistency and good-to-excellent inter-rater reliability. The authors conclude that this tool can serve as a measure of the integrity of BI delivery in randomized clinical trials and as an integral component of BI skills-based teaching exercises allowing trainers to provide structured feedback to trainees. Further research is needed in more diverse populations to understand its application in practice. However, as the field of integrated medical and addiction care matures, standardized assessments will be essential to efforts to improve the quality of brief interventions.

2.2. Quality of chronic disease management for alcohol and other drug treatment in primary care

Although alcohol and other drug (AOD) dependence are increasingly recognized as chronic conditions, treatment of these disorders often remains stuck in an acute-care paradigm (Friedmann, Lemon, Stein, & D'Aunno, 2003). Implementing a chronic disease management approach to treatment of AOD dependence has the potential to improve addiction outcomes and patient satisfaction by providing flexible, patient-centered, longitudinal care (Samet et al., 2001; McKay, 2009). Support for chronic disease management (CDM) models for addiction treatment exists; however, no formal assessment exists to measure the quality of CDM for AOD. Visit frequency and timing-based measures have classically been considered quality assessments in this area (McCorry, Garnick, Bartlett, Cotter, & Chalk, 2000), whereas consideration of addiction outcomes has less commonly been considered. As such, Kim et al., 2012 investigated the relationship between quality of CDM measures for AOD dependence and addiction outcomes using data from a randomized clinical trial of access to primary care-based CDM for AOD dependence. Visit frequency, based on engagement in the CDM clinic, was not associated with abstinence or addiction severity while self-reported perception of receipt of chronic disease care, regardless of the source of care, was associated with abstinence and alcohol addiction severity. This work supports the efforts to integrate primary health care, mental health services and addiction treatment while showing that the quality of the CDM provided needs to be considered. Specifically, as we continue to understand and implement integrated medical systems, patient perception of receipt of quality CDM may be a key factor associated with improved disease-specific outcomes.

2.3. The potential of technology to improve the delivery of screening and brief intervention in medical settings

Several barriers exist to implementing behavioral health screening and interventions in emergency department settings, such as time, provider constraints, lack of provider training, and incomplete knowledge of area referral and follow-up resources (Delgado, Acosta, Ginde, et al., 2011). Using computer interfaces and other technology to deliver substance use screening and intervention has been proposed to perform these services, bypassing individual practitioner time and expertise and allowing modifiable messages based on patient responses. Choo et al, 2012. explored the accessibility and interest in technology-based substance use health information among patients with alcohol and drug use presenting to the emergency department and the barriers to using such technology. The results of this study suggest that although access to technology was high, interest in receiving health information via technology ranged from 41 to 46% among patients misusing drugs and alcohol. Barriers included concerns over confidentiality—especially among patients with drug misuse, complexity of use and time. Technology will undoubtedly play a key role in the implementation of integrated health care delivery systems. It is clear, however, that confidentiality concerns must be addressed if integrated health care settings are to realize the full potential of technology for screening and intervention.

2.4. Quality of behavioral health integration in medical care

Health care reform has provided much impetus for comprehensive and integrated care, including integration of mental health and substance use services into primary care settings (Humphreys & McLellan, 2011). Federally Qualified Health Centers (FQHCs) will play a central role in these efforts, especially for patients with substance use disorders. McGovern et al., 2012 describe the development and feasibility of an organizational measure, the Dual Diagnosis Capability in Health Care Settings (DDCHCS) index, to assess the level of behavioral health integration in FQHCs. Prior work has established the utility of the Dual Diagnosis Capability in Addiction Treatment and the Dual Diagnosis in Mental Health Treatment to inform policy makers and researchers in efforts toward integrated care delivery (McGovern, Lambert-Harris, McHugo, Giard, & Mangrum, 2010). After establishing DDCHCS feasibility and reliability, the authors utilized this measure to assess 13 FQHCs on level of behavioral health and substance use integration. They found that 3 of 13 FQHCs met standards to be dual diagnosis capable and were more likely to be integrated in behavioral health rather than addiction treatment services. Significant variation was noted in center staffing, treatment practices and program milieu. Although designed to assess feasibility and field testing of the DDCHCS, this study suggests that further work is needed to enhance the capability of FQHCs to deliver integrated health care to the large population of patients with substance use disorders who will become insured with the Medicaid expansion under the Affordable Care Act.

3. Meeting the challenge of substance use and co-occurring disorders

3.1. Screening and brief intervention for alcohol use uncovers drug and mental health issues

The American College of Surgeons mandates that trauma centers have alcohol screening and brief intervention services integrated into their programs. However, screening for alcohol use may uncover other problems, such as problem drug use and mental illness. Among a cohort of 878 injured trauma survivors, Zatzick et al., 2012 identify a high rate (74%) of illegal drug or alcohol problems and 25% with post-traumatic stress disorder (PTSD). Among the 166 (19%) patients who received the brief intervention for alcohol, 70% had a concomitant illegal drug use or PTSD. So, while alcohol screening and brief intervention are mandated, it uncovers substantial other substance use and mental health issues, which these programs do not typically address. This finding demonstrates that integration to improve quality and efficiency can yield further challenges in the real world. The authors suggest stepped-care interventions as a potential solution,

where the intervention adapts based on the individual's co-morbidities. Whether stepped-care interventions will work awaits further study.

3.2. Integrated models facilitate treatment of HIV and hepatitis C in addiction treatment settings

Methadone maintenance treatment for opioid dependence in the United States requires daily face-to-face dosing at the beginning of treatment, with the possibility of earning take-home doses up to once every 2 weeks. This regular patient–provider contact creates the opportunity to deliver observed medical treatment for co-morbid medical conditions. Two articles in this issue describe methadone clinic-based treatment programs that seek to improve adherence to medication treatment, one offering directly observed HIV medication dosing, the other offering concurrent group treatment for hepatitis C infection. In both models, patients and providers chose to extend the duration of the integrated care intervention beyond the study period to cover the course of medical treatment. Sorensen et al. report on the implementation of a directly administered anti-retroviral therapy program among 24 HIV-infected methadone patients. They found improved HIV viral suppression and high rates of adherence to HIV treatment. Initially intended as a 24-week pilot that would include a planned step-down in intensity, at 48 weeks 38% of the patients were still receiving daily directly administered treatment because the treatment providers and patients were not willing to transition the patients to lower intensity. This small study suggests that directly administered HIV treatment in methadone programs is feasible, popular among patients and staff, and improves HIV outcomes.

Stein et al., 2012 describe the novel implementation of a concurrent hepatitis C infection group treatment program among 42 opioid dependent patients, most of whom were receiving methadone maintenance at multi-disciplinary drug treatment center. Concurrent group treatment included weekly group meetings for patients who initiated and continued through treatment together, experiencing the same treatment milestones together. Groups were co-led by a medical provider and peer educator and included elements of a traditional medical visit such as checking vital signs, administering pegylated interferon injections and provision of prescriptions during the group visit. This treatment program achieved treatment success rates comparable to previously published randomized clinical trials (Jacobson, McHutchison, Dusheiko, et al., 2011; Litwin, Harris, Nahvi, et al., 2009; Poordad, McCone, Bacon, et al., 2011; Rumi, Aghemo, Prati, et al., 2010; von Wagner, Huber, Berg, et al., 2005). This program, like that of Sorensen et al. above, was extended beyond the initially-intended 12 weeks to the full treatment duration (up to 48 weeks) because almost all patients wanted to continue group treatment. This study shows the promise of group medical treatment for hepatitis C within addiction treatment settings. It merits further study in controlled trials with more diverse addiction treatment populations —including those on buprenorphine maintenance treatment.

3.3. Primary care models can integrate addiction and medical care for HIV and hepatitis C

Office-based opioid dependence treatment with buprenorphine is effective in primary care settings (Fiellin, Pantalon, Chawarski, et al., 2006). Clinical guidelines mandate that counseling be available, yet how much counseling should delivered and by whom is not clear (Sullivan, Barry, Moore, et al., 2006). Tetrault et al., 2012 report the results of a randomized controlled trial of 15-minute physician-delivered counseling (PM) every other week compared to PM plus an additional 45-minute nurse-delivered counseling every week for opioid dependent patients initiating buprenorphine/naloxone treatment in an HIV primary care clinic. No differences in addiction or HIV-related outcomes were demonstrated. However, patients in both groups had improved control of their HIV when taking buprenorphine/naloxone. This study provides further confirmation that buprenorphine/naloxone treatment integrated into an HIV clinic has multiple benefits for opioid-dependent, HIV-infected patients, regardless of the level of counseling available.

For treatment of hepatitis C infection, achieving the success rates found in clinical trials in less selected, more diverse urban populations has been challenging (Feuerstadt, Bunim, Garcia, et al., 2010). Initiation of hepatitis C treatment in real-world patients is a challenge because of the subacute presentation, competing priorities, numerous relative contraindications, and the length, complexity, and side effects of the medication. Islam et al., 2012. describe an Australian effort to increase uptake of hepatitis C treatment through screening and facilitated referral from an integrated primary care and syringe access program. Of 143 patients with confirmed hepatitis C infection, 67% were referred to specialty care and 48% attended at least one appointment, but only 7.5% initiated hepatitis C treatment. This study shows that a primary care model integrated into a syringe access program can facilitate an appointment with a specialist, but treatment initiation remains low. In order to treat more of those at highest risk for hepatitis C infection—people who inject drugs—innovative and integrated programs are needed that build on the experience of Islam et al., 2012.

3.4. Buprenorphine/naloxone for co-occurring chronic pain and opioid dependence

Caring for patients with chronic pain and opioid dependence is challenging because opioid treatment both reduces the pain and has the potential to worsen addiction (Barry, Irwin, Jones, et al., 2010). Pade et al., 2012 describe the concomitant treatment of chronic pain and opioid dependence with buprenorphine/naloxone within a primary care-based co-occurring disorders clinic at the Albuquerque Veteran Affairs Medical Center. The clinic provided team-based primary care, opioid dependence treatment and psychiatric evaluation for patients with chronic pain and opioid dependence. Patient retention was high and pain scores improved in pre–post analyses. While chronic pain is a relative contraindication for buprenorphine treatment because of its partial agonist ceiling effect, this study supports the use of buprenorphine/naloxone by a primary care-based multi-disciplinary care team as a well-tolerated alternative to full opioid agonists for many patients with chronic pain and opioid dependence. Further research should focus on buprenorphine dose ranging and standardization of pain outcomes and may ultimately determine whether buprenorphine/naloxone should be considered as a first-line treatment for patients with chronic pain at high risk for opioid dependence and/or overdose.

4. Realizing the potential of addiction pharmacotherapy

When medications are prescribed for substance use disorders, these conditions unavoidably become medicalized. The people with substance use disorders become patients, as a prescriber now must be involved in the care. Side effects and drug interactions need to be monitored. The study by Pade et al., 2012 described above demonstrated how integration of treatment for both medical and substance use conditions occurs by co-locating the care, but also can hinge on the innovative use of medications to treat two problems at the same time. Two other studies in this issue describe pharmacotherapy innovations within integrated treatment models.

4.1. Buprenorphine/naloxone by mobile medical van

Schwarz et al., 2012 describe the delivery of buprenorphine/naloxone treatment for opioid dependence among a population that was not been included in prior randomized controlled trials. They report on Project BEST (Buprenorphine Entry into Substance Abuse Treatment), a mobile medical van-based program linked to a syringe

access program that brought medical care and buprenorphine/naloxone treatment for opioid dependence into four impoverished neighborhoods in New Haven, CT. Patients retained on buprenorphine/naloxone for more than 1 week had fewer emergency department visits that those retained less than 1 week, but hospitalizations and lengths of stay did not differ. Although not a conclusive trial, this study demonstrates that retaining patients in community-based, medication-oriented addiction care might reduce utilization of emergency care.

4.2. Extended-release naltrexone for alcohol dependence delivered in primary care

Extended-release naltrexone (XR-NTX) is FDA-approved for the treatment of alcohol dependence and opioid dependence. It is typically dosed monthly via a medical provider-delivered injection. Lee et al. (2010) previously demonstrated a 56% treatment completion rate and improved drinking outcomes at 12 weeks using XR-NTX in a primary care clinic with physician-delivered medical management counseling. The current study describes extended follow-up (Lee). Of 40 patients who completed the initial 12-week study, 19 (48%) continued treatment for a median of 38 weeks with persistently low levels of drinking. Of the original 72 patients who entered the original study, 7 (10%) completed the full 15 months of treatment. Thus, in this real world primary care-based study, a small group did continue to benefit from long-term XR-NTX. Further research should focus on improving retention and developing alternative strategies for those patients who are less successful.

5. Conclusions

As the implementation of health care reform progresses with conceptual support for the integration of the screening and treatment of substance use disorders and medical conditions, this special issue of the *Journal of Substance Abuse Treatment* offers multiple examples of innovative care models that put the concept of integration into practice. These papers examine important issues underlying the quality of substance use intervention delivery in medical settings; detection and co-management of co-morbidities that may be uncovered when integrated care is delivered; and innovative pharmacotherapy strategies. Integration is a key feature of health system reform that promises to improve the access, quality and value of care for all patients, but especially those with co-occurring substance use, mental health and medical disorders. Hopefully, the insights from this collection will inform this important work.

References

Academic, E. D. (2007). SBIRT Research Collaborative. The impact of screening, brief intervention, and referral for treatment on emergency department patients' alcohol use. *Annals of Emergency Medicine*, *50*, 699–710.

Babor, T. F., McRee, B. G., Kassebaum, P. A., Grimaldi, P. L., Ahmed, K., & Bray, J. (2007). Screening, Brief Intervention, and Referral to Treatment (SBIRT): Toward a public health approach to the management of substance abuse. *Substance Abuse*, *28*, 7–30.

Barry, D. T., Irwin, K. S., Jones, E. S., et al. (2010). Opioids, chronic pain, and addiction in primary care. *The Journal of Pain*, *11*, 1442–1450.

Bernstein, E., Bernstein, J., Feldman, J., et al. (2007). An evidence based alcohol Screening, Brief Intervention and Referral to Treatment (SBIRT) curriculum for emergency department (ED) providers improves skills and utilization. *Substance Abuse*, *28*, 79–92.

Choo, E. K., Ranney, M. L., Wong, Z., & Mello, M. J. (2012). Attitudes toward technology-based health information among adult emergency department patients with drug or alcohol misuse. *Journal of Substance Abuse Treatment*, *43*, 397–401.

Committee on Crossing the Quality Chasm. (2006). Adaptation to Mental Health and Addictive Disorders. Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series. (Accessed 9/6/2012 at http://wwwnapedu/catalogphp?record_id=11470#toc). Washington, DC: National Academies Press.

Delgado, M. K., Acosta, C. D., Ginde, A. A., et al. (2011). National survey of preventive health services in US emergency departments. *Annals of Emergency Medicine*, *57*, 104–108.e2.

D'Onofrio, G., Nadel, E. S., Degutis, L. C., et al. (2002). Improving emergency medicine residents' approach to patients with alcohol problems: A controlled educational trial. *Annals of Emergency Medicine*, *40*, 50–62.

D'Onofrio, G., Pantalon, M. V., Degutis, L. C., et al. (2008). Brief intervention for hazardous and harmful drinkers in the emergency department. *Annals of Emergency Medicine*, *51*, 742 .e2–750 .e2.

Estee, S., Wickizer, T., He, L., Shah, M. F., & Mancuso, D. (2010). Evaluation of the Washington state screening, brief intervention, and referral to treatment project: Cost outcomes for Medicaid patients screened in hospital emergency departments. *Medical Care*, *48*, 18–24.

Feuerstadt, P., Bunim, A. L., Garcia, H., et al. (2010). Effectiveness of hepatitis C treatment with pegylated interferon and ribavirin in urban minority patients. *Hepatology*, *51*, 1137–1143.

Fiellin, D. A., Pantalon, M. V., Chawarski, M. C., et al. (2006). Counseling plus buprenorphine–naloxone maintenance therapy for opioid dependence. *The New England Journal of Medicine*, *355*, 365–374.

Fleming, M. F., Barry, K. L., Manwell, L. B., Johnson, K., & London, R. (1997). Brief physician advice for problem alcohol drinkers. A randomized controlled trial in community-based primary care practices. *JAMA : The Journal of the American Medical Association*, *277*, 1039–1045.

Friedmann, P. D., Lemon, S. C., Stein, M. D., & D'Aunno, T. A. (2003). Accessibility of addiction treatment: Results from a national survey of outpatient substance abuse treatment organizations. *Health Services Research*, *38*, 887–903.

Gentilello, L. M., Ebel, B. E., Wickizer, T. M., Salkever, D. S., & Rivara, F. P. (2005). Alcohol interventions for trauma patients treated in emergency departments and hospitals: A cost benefit analysis. *Annals of Surgery*, *241*, 541–550.

Havard, A., Shakeshaft, A., & Sanson-Fisher, R. (2008). Systematic review and meta-analyses of strategies targeting alcohol problems in emergency departments: Interventions reduce alcohol-related injuries. *Addiction*, *103*, 368–376 [discussion 77–8].

Humphreys, K., & McLellan, A. T. (2011). A policy-oriented review of strategies for improving the outcomes of services for substance use disorder patients. *Addiction*, *106*, 2058–2066.

Islam, M. M., Topp, L., Conigrave, K. M., White, A., Reid, S. E., Grummett, S., Haber, P. S., & Day, C. A. (2012). Linkage into specialist hepatitis C treatment services of injecting drug users attending a needle syringe program-based primary healthcare centre. *Journal of Substance Abuse Treatment*, *43*, 440–445.

Jacobson, I. M., McHutchison, J. G., Dusheiko, G., et al. (2011). Telaprevir for previously untreated chronic hepatitis C virus infection. *The New England Journal of Medicine*, *364*, 2405–2416.

Kim, T. W., Saitz, R., Cheng, D. M., Winter, M. R., Witas, J., & Samet, J. H. (2012). Effect of quality chronic disease management for alcohol and drug dependence on addiction outcomes. *Journal of Substance Abuse Treatment*, *43*, 389–396.

Lee, J. D., Grossman, E., DiRocco, D., et al. (2010). Extended-release naltrexone for treatment of alcohol dependence in primary care. *Journal of Substance Abuse Treatment*, *39*, 14–21.

Lee, J. D., Grossman, E., Huben, L., Manseau, M., McNeely, J., Rotrosen, J., Stevens, D., & Gourevitch, M. N. (2012). Extended-release naltrexone plus medical management alcohol treatment in primary care: findings at 15 months. *Journal of Substance Abuse Treatment*, *43*, 458–462.

Litwin, A. H., Harris, K. A., Jr., Nahvi, S., et al. (2009). Successful treatment of chronic hepatitis C with pegylated interferon in combination with ribavirin in a methadone maintenance treatment program. *Journal of Substance Abuse Treatment*, *37*, 32–40.

McCorry, F., Garnick, D. W., Bartlett, J., Cotter, F., & Chalk, M. (2000). Developing performance measures for alcohol and other drug services in managed care plans. Washington Circle Group. *The Joint Commission Journal on Quality Improvement*, *26*, 633–643.

McGovern, M. P., Lambert-Harris, C., McHugo, G. J., Giard, J., & Mangrum, L. (2010). Improving the dual diagnosis capability of addiction and mental health treatment services: Implementation factors associated with program level changes. *J Dual Diagnosis*, *6*, 237–250.

McGovern, M. P., Urada, D., Lambert-Harris, C., Sullivan, S. T., & Mazade, N. A. (2012). Development and initial feasibility of an organizational measure of behavioral health integration in medical care settings. *Journal of Substance Abuse Treatment*, *43*, 402–409.

McKay, J. R. (2009). Continuing care research: What we have learned and where we are going. *Journal of Substance Abuse Treatment*, *36*, 131–145.

Pade, P. A., Cardon, K. E., Hoffman, R. M., & Geppert, C. M. A. (2012). Prescription opioid abuse, chronic pain, and primary care: A Co-Occurring Disorders Clinic in the chronic disease model. *J Subst Abuse Treat*, *43*, 446–450.

Pantalon, M. V., Martino, S., Dziura, J., Li, F. -Y., Owens, P. H., Fiellin, D. A., O'Connor, P. G., & D'Onofrio, G. (2012). Development of a scale to measure practitioner adherence to a brief intervention in the emergency department. *Journal of Substance Abuse Treatment*, *43*, 382–388.

Poordad, F., McCone, J., Jr., Bacon, B. R., et al. (2011). Boceprevir for untreated chronic HCV genotype 1 infection. *The New England Journal of Medicine*, *364*, 1195–1206.

Rumi, M. G., Aghemo, A., Prati, G. M., et al. (2010). Randomized study of peginterferon-alpha2a plus ribavirin vs peginterferon-alpha2b plus ribavirin in chronic hepatitis C. *Gastroenterology*, *138*, 108–115.

Ryan, S. A., Martel, S., Pantalon, M., et al. (2012). Screening, Brief Intervention, and Referral to Treatment (SBIRT) for alcohol and other drug use among adolescents: Evaluation of a pediatric residency curriculum. *Substance Abuse*, *33*, 251–260.

Samet, J. H., Friedmann, P., & Saitz, R. (2001). Benefits of linking primary medical care and substance abuse services: Patient, provider, and societal perspectives. *Archives of Internal Medicine*, *161*, 85–91.

Schwarz, R., Zelenev, A., Bruce, R. D., & Altice, F. L. (2012). Retention on buprenorphine treatment reduces emergency department utilization, but not hospitalization, among treatment-seeking patients with opioid dependence. *Journal of Substance Abuse Treatment*, *43*, 451–457.

Sorensen, J. L., Haug, N. A., Larios, S., Gruber, V. A., Tulsky, J., Powelson, E., Logan, D. P., & Shapiro, B. (2012). Directly administered antiretroviral therapy: Pilot study of a structural intervention in methadone maintenance. *Journal of Substance Abuse Treatment*, *43*, 418–423.

Stein, M. R., Soloway, I. J., Jefferson, K. S., Roose, R. J., Arnsten, J. H., & Litwin, A. H. (2012). Concurrent group treatment for hepatitis C: Implementation and outcomes in a methadone maintenance treatment program. *Journal of Substance Abuse Treatment*, *43*, 424–432.

Substance Abuse and Mental Health Services Administration. (2012). Screening, Brief Intervention and Referral to Treatment (SBIRT). (Accessed 8/8/2012, at http://www.samhsa.gov/prevention/sbirt/).

Sullivan, L. E., Barry, D., Moore, B. A., et al. (2006). A trial of integrated buprenorphine/naloxone and HIV clinical care. *Clinical Infectious Diseases*, *43*(Suppl 4), S184–S190.

Tetrault, J. M., Moore, B. A., Barry, D. T., O'Connor, P. G., Schottenfeld, R., Fiellin, D. A., & Fiellin, L. E. (2012). Brief versus extended counseling along with buprenorphine/naloxone for HIV-infected opioid dependent patients. *Journal of Substance Abuse Treatment*, *43*, 433–439.

Tetrault, J. M., Green, M. L., Martino, S., et al. (2012). Developing and implementing a multispecialty graduate medical education curriculum on Screening, Brief Intervention, and Referral to Treatment (SBIRT). *Substance Abuse*, *33*, 168–181.

von Wagner, M., Huber, M., Berg, T., et al. (2005). Peginterferon-alpha-2a (40KD) and ribavirin for 16 or 24 weeks in patients with genotype 2 or 3 chronic hepatitis C. *Gastroenterology*, *129*, 522–527.

Zatzick, D., Donovan, D., Dunn, C., Russo, J., Wang, J., Jurkovich, G., Rivara, F., Whiteside, L., Ries, R., & Gentilello, L. (2012). Substance use and posttraumatic stress disorder symptoms in trauma center patients receiving mandated alcohol screening and brief intervention. *Journal of Substance Abuse Treatment*, *43*, 410–417.

Genetic, Physiological, and Lifestyle Predictors of Mortality in the General Population

Stefan Walter, MA, MSc, MPH, Johan Mackenbach, MD, PhD, Zoltán Vokó, MD, PhD, Stefan Lhachimi, MPP, MSc, M. Arfan Ikram, MD, PhD, André G. Uitterlinden, PhD, Anne B. Newman, MD, MPH, Joanne M. Murabito, MD, ScM, Melissa E. Garcia, MPH, Vilmundur Gudnason, MD, PhD, Toshiko Tanaka, PhD, Gregory J. Tranah, PhD, Henri Wallaschofski, MD, Thomas Kocher, MD, Lenore J. Launer, PhD, Nora Franceschini, MD, MPH, Maarten Schipper, PhD, Albert Hofman, MD, PhD, and Henning Tiemeier, MD, PhD

Objectives. We investigated the quality of 162 variables, focusing on the contribution of genetic markers, used solely or in combination with other characteristics, when predicting mortality.

Methods. In 5974 participants from the Rotterdam Study, followed for a median of 15.1 years, 7 groups of factors including age and gender, genetics, socioeconomics, lifestyle, physiological characteristics, prevalent diseases, and indicators of general health were related to all-cause mortality. Genetic variables were identified from 8 genome-wide association scans (n = 19 033) and literature review.

Results. We observed 3174 deaths during follow-up. The fully adjusted model (C-statistic for 15-year follow-up [C_{15y}] = 0.80; 95% confidence interval [CI] = 0.75, 0.77) predicted mortality well. Most of the additional information apart from age and sex stemmed from physiological markers, prevalent diseases, and general health. Socioeconomic factors and lifestyle contributed meaningfully to mortality risk prediction with longer prediction horizon. Although specific genetic factors were independently associated with mortality, jointly they contributed little to mortality prediction (C_{15y} = 0.56; 95% CI = 0.55, 0.57).

Conclusions. Mortality can be predicted reasonably well over a long period. Genetic factors independently predict mortality, but only modestly more than other risk indicators. (*Am J Public Health.* 2012;102:e3–e10. doi:10.2105/AJPH.2011.300596)

In the 20th century, life expectancy at birth increased from 50 years to over 80 years in Western countries.[1] Demographers repeatedly predicted that it had reached a ceiling, but life expectancy in record countries continues to rise by an average of 3 months each year.[2] Although epidemiological research has identified numerous predictors of mortality, information about their comparative effect sizes and long-term predictive power is sparse. Prior research has often been limited by a short period of follow-up, a limited set of covariates, or a focus on cause-specific mortality. Only a few studies have evaluated the potential for explaining mortality from a broader perspective by jointly analyzing demographic characteristics, lifestyle factors, and indicators of health and disease.[3–7] It is still unclear whether genetic information can be used to predict mortality, but recent advances in genomic technology allow for the inclusion of genetic markers in the prediction of mortality.

We combined traditional indicators of mortality risk with genetic factors, derived from a meta-analysis of 8 genome-wide association studies and the literature, and associated them with mortality over 15 years of follow-up. Our aims were twofold: first, to identify independent determinants of mortality by analyzing 162 a priori identified risk factors; second, to provide information on the independent and combined potential of genetic markers in predicting mortality.

METHODS

The Rotterdam Study is a population-based prospective cohort initiated in 1990. It was designed to investigate risk factors for diseases in 7983 participants aged 55 years or older in the Ommoord district of Rotterdam, The Netherlands.[8,9] In the initial and subsequent investigation waves, trained research assistants collected data on health, medication use, medical and family history, and lifestyle factors in extensive home interviews. Participants subsequently visited the research center for clinical examinations, with a special emphasis on imaging and on collecting and storing biospecimens to facilitate in-depth molecular and genetic analyses. Data analyzed in this study concern 5974 participants with genetic information available from the first wave of the Rotterdam Study.

Predictors

We organized baseline data into related groups: age and gender, genetics, socioeconomics, lifestyle, physiology, diseases, and general health. We hypothesized a priori that genetics, socioeconomics, and lifestyle were associated with long-term health effects, whereas physiology, disease, and general health were more likely associated with short-term mortality.[10] Overall, we analyzed 162 risk indicators in this study: 69 previously studied risk factors for mortality and 93 single nucleotide polymorphisms (SNPs).

To study possible genetic risk factors of mortality, we used the genetic data of 19 033 participants (women = 55%) aged 55 years and older from 8 discovery cohorts of European ancestry. These people were participants not of the Rotterdam Study but of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE), of which the Rotterdam Study is a part.[11] We identified the top 50 loci from the meta-analysis of genome-wide association studies on time to death. In addition, we used 43 SNPs that mapped to genes from a seminal review.[1] For the analysis, we extracted all SNPs from the imputed gene information of the Rotterdam Study, except for apolipoprotein E (*APOE*), which we genotyped directly.[12]

TABLE 1—Descriptive Statistics (Unadjusted and Completely Adjusted Models) and Association to 15-Year Mortality: The Rotterdam Study, 1990–2009

Variable	Baseline, Mean ±SD or No. (%)	Univariate		Final Model	
		RR (95% CI)	*P*	RR (95% CI)	*P*
Age, y	69.43 ±9.10	1.12 (1.11, 1.12)	< .001	1.09 (1.08, 1.10)	< . 001
Female gender	3547 ±59.38	0.79 (0.73, 0.85)	< .001	0.71 (0.62, 0.81)	< . 001
Genetics (rs, gene, Chr, allele 1/2)[a]					
Candidate genes from literature					
APOE ε 4 allele, chr 19	0.28	1.01 (0.95, 1.08)	.71	1.10 (1.03, 1.19)	.007
rs6997892, *WRN*, chr 8 (G/A)	0.88	0.97 (0.91, 1.05)	.48	0.92 (0.86, 1.00)	.04
rs2684766, *IGF1R*, chr 15 (T/C)	0.97	1.03 (0.88, 1.19)	.75	1.19 (1.00, 1.40)	.05
rs11630259, *IGF1R*, chr 15 (T/C)	0.73	1.04 (0.99, 1.10)	.15	1.09 (1.02, 1.17)	.02
GWA continuous mortality selection					
rs10817931, *TRIM32*, chr 9(A/C)	0.38	1.02 (0.97, 1.07)	.54	1.07 (1.01, 1.13)	.01
rs1421783, *MAT2B*, chr 5(C/G)	0.93	0.94 (0.85, 1.04)	.24	0.89 (0.80, 1.00)	.05
Socioeconomic and lifestyle characteristics					
Social class (min = 1, max = 4)	2.59 ±1.19	0.87 (0.84, 0.90)	< .001	0.95 (0.91, 0.99)	.03
Living situation			< .001		.05
Independent	4941 (82.71)	1.00 (Ref)		1.00 (Ref)	
Service flat	610 (10.21)	2.88 (2.61, 3.18)		1.04 (0.93, 1.16)	
Home for elderly	423 (7.08)	9.06 (8.12, 10.12)		1.26 (1.03, 1.47)	
Smoking			.53		< .001
Never	2101 (35.17)	1.00 (Ref)		1.00 (Ref)	
Former	2491 (41.70)	0.94 (0.87, 1.02)		1.07 (0.96, 1.19)	
Current	1382 (23.13)	1.12 (1.02, 1.23)		1.45 (1.27, 1.66)	
Pack-years	16.58 ±23.15	1.12 (1.08, 1.16)	< .001	1.07 (1.03, 1.12)	< .001
Nutrition: energy intake, kJ	8280.21 ± 2133	1.05 (0.99, 1.11)	.11	1.08 (1.02, 1.13)	.006
Physiological characteristics					
Diastolic blood pressure, mmHg	73.71 ±11.50	0.99 (0.95, 1.02)	.49	1.05 (1.00, 1.11)	.04
Systolic blood pressure, mmHg	139.37 ±22.30	1.37 (1.32, 1.41)	< .001	1.06 (1.00, 1.13)	.03
Body mass index, kg/m^2	26.30 ±3.71	0.93 (0.89, 0.97)	.05	0.86 (0.80, 0.92)	< .001
Body mass index squared, $(kg/m^2)^2$	705.33 ±205.78	1.04 (1.02, 1.06)	.05	1.03 (1.01, 1.05)	< .001
Waist circumference, cm (SD)	90.57 ±11.17	1.21 (1.16, 1.25)	< .001	1.10 (1.04, 1.17)	.002
Erythrocyte sedimentation, mm/h	13.53 ±11.89	1.32 (1.28, 1.37)	< .001	1.08 (1.02, 1.14)	.006
Leukocytes, $\times 10^9/L$	6.70 ±1.92	1.18 (1.15, 1.21)	< .001	1.11 (1.07, 1.15)	< .001
Creatinine, μmol/L	83.18 ±20.52	1.20 (1.18, 1.22)	< .001	1.06 (1.01, 1.12)	.03
C-reactive protein	3.36 ±6.61	1.21 (1.19, 1.23)	< .001	1.07 (1.03, 1.10)	< .001
Total cholesterol, mmol/L	6.60 ±1.22	0.81 (0.78, 0.84)	< .001	0.92 (0.89, 0.96)	< .001
Bone mineral density of femoral neck	0.83 ±0.14	0.77 (0.71, 0.83)	< .001	0.93 (0.88, 0.99)	.01
Aortic calcification	1.80 ±1.49	1660 (1.55, 1.76)	< .001	1.08 (1.01, 1.16)	.03
Disease characteristics					
Diabetes mellitus (yes vs no)	618 (10.35)	2.11 (1.92, 2.33)	< .001	1.39 (1.25, 1.55)	< .001
Left ventricular hypertrophy (yes vs no)	258 (4.32)	2.35 (2.04, 2.70)	< .001	1.33 (1.13, 1.55)	< .001
Atrial fibrillation (yes vs no)	318 (5.32)	3.28 (2.89, 3.73)	< .001	1.32 (1.15, 1.51)	< .001
Peripheral artery disease (yes vs no)	1133 (18.97)	2.63 (2.42, 2.86)	< .001	1.16 (1.03, 1.30)	.01
Myocardial infarction (yes vs no)	754 (12.62)	2.06 (1.87, 2.26)	< .001	1.39 (1.25, 1.55)	< .001
Disease (yes vs no)	64 (1.07)	4.15 (3.26, 5.28)	< .001	1.54 (1.16, 2.05)	.003

Continued

TABLE 1—*Continued*

Prevalent cancer			< .001		< .001
Time, 0–5 y (yes vs no)	282 (4.72)	2.58 (2.05, 3.24)		2.03 (1.60, 2.58)	
Time, > 5 y (yes vs no)	200 (3.87)	1.44 (1.18, 1.76)		1.08 (0.88, 1.30)	
Mini-Mental State Examination	27.26 ±2.84	0.59 (0.58, 0.61)	< .001	0.86 (0.82, 0.90)	< .001
General health					
Serious illness in the last 5 y? (yes vs no)	621 (10.40)	1.48 (1.33, 1.66)	< .001	1.13 (1.00, 1.28)	.05
Unintentional weight loss? (yes vs no)	675 (11.30)	2.12 (1.92, 2.33)	< .001	1.22 (1.09, 1.36)	< .001
How is your general health compared with members of your age group?			< .001		< .001
Better	3083 (51.61)	1.00 (Ref)		1.00 (Ref)	
Same	2299 (38.48)	0.97 (0.90, 1.05)		1.06 (0.97, 1.15)	
Worse	592 (9.91)	1.59 (1.42, 1.77)		1.32 (1.14, 1.53)	
Prevalent memory complaints					
Time, 0–5 y			< .001		.003
No memory complaints	5559 (93.06)	1.00 (Ref)		1.00 (Ref)	
Mild memory complaints	370 (6.19)	4.86 (4.05, 5.82)		1.15 (0.94, 1.40)	
Severe memory complaints	45 (0.75)	10.36 (7.31, 14.67)		1.02 (0.64, 1.63)	
Time, > 5 y			< .001		.003
No memory complaints	4928 (95.41)	1.00 (Ref)		1.00 (Ref)	
Mild memory complaints	207 (4.01)	2.61 (2.19, 3.1)		1.07 (0.90, 1.27)	
Severe memory complaints	15 (0.29)	17.04 (10.25, 28.33)		3.34 (1.82, 6.11)	

Note. Chr = chromosome; CI = confidence interval; CVA = cardiovascular accident; GWA = genome-wide association; RR = relative risk. Total number of participants was 5974. The table shows all variables significant in the final model. Variables included in the full model but not included in this table are the following: socioeconomics: education, social class, occupation, insurance, living circumstance, death of spouse; lifestyle: alcohol consumption in g/day, fruit consumption in g/day, vegetable consumption in g/day; physiology: waist circumference, high density lipoprotein cholesterol, bone mineral density of lumbar spine, bone mineral density of femoral neck; general health: specialist visit within the last month, number of specialist visits in the last year, general practitioner visit within the last month, number of general practitioner visits in the last year, hospitalization within the last year, falls in the previous month, activities of daily living; disease: gout, vertebral fracture, cardiovascular accident, transient ischemic attack, hip fracture, coronary operation.
[a]Values represent the frequency of allele 1.

Information about cohort characteristics, genotyping, and imputation of the discovery set is presented in Tables A and B (available as a supplement to the online version of this article at http://www.ajph.org).

We included the following variables describing the socioeconomic status of the study population: education, employment status, monthly income, social class (derived from occupation of head of the household), health insurance status, number of children, whether living independently or in a nursing home, whether living with a partner, and death of spouse.[8,9] We included the following indicators of lifestyle: riding a bike, alcohol consumption, smoking, energy intake, and fruit and vegetable consumption.[13,14]

We assessed physiological characteristics in terms of body weight, body mass index, waist circumference, hip circumference, waist-to-hip ratio, sitting systolic and diastolic blood pressures, leukocyte count, erythrocyte sedimentation rate, albumin level, total cholesterol level, high-density lipoprotein cholesterol, creatinine, uric acid, serum C-reactive protein, postload insulin, bone mineral density of the femoral neck and lumbar spine, and atherosclerotic plaques. We assessed all physiological variables using standard medical, laboratory, and imaging procedures as described previously.[15–20]

On the basis of self-report, investigations at the baseline center visit, medical record information, and drug utilization, we defined the following prevalent diseases: diabetes, left ventricular hypertrophy, atrial fibrillation, hypertension, hip fracture, peripheral artery disease, myocardial infarction, heart failure, dementia, gout, Parkinson's disease, stroke, transient ischemic attack, cancer, cognitive function (Mini-Mental State Examination), and coronary operation.[21–27]

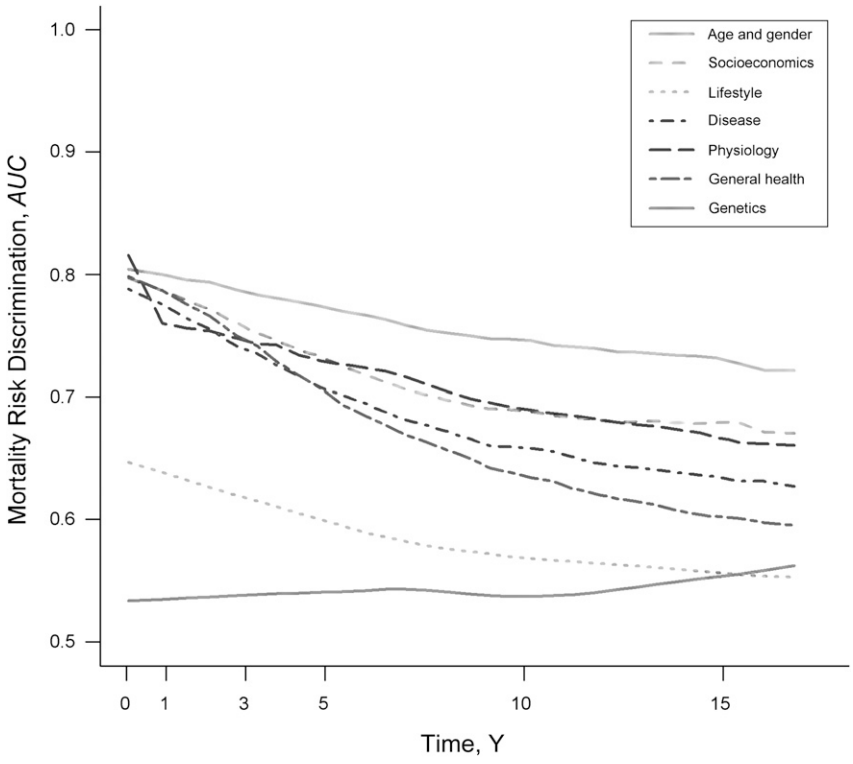
General health included the following: activities of daily living[28] and instrumental activities of daily living,[29] health care utilization, self-perceived comparative health, accidental falls, shortness of breath, past serious illness and hospitalization, unintentional weight loss, and self-reported memory complaints.

Outcome

The outcome measure was time to death from any cause. All participants of the Rotterdam Study were under continuous surveillance; general practitioners' and hospital records, as well as death certificates from the municipality, were used to identify participants who died before January 1, 2009. The median follow-up was 15.1 years (range = 0.05–19.50).

Statistical Analysis

We used SAS 9.2 (SAS Institute Inc, Cary, NC) and the PROC MI procedure to impute to 5 complete data sets. We set the maximum missingness for analysis of the data at 30% a priori. The percentage of missing information is reported in Table C (available as a supplement



Note. AUC = area under the curve.

FIGURE 1—Time-dependent receiver-operating characteristic curves for prediction of mortality for different groups of variables: The Rotterdam Study, 1990–2009.

to the online version of this article at http://www.ajph.org). Other than age, all continuous variables were standardized to facilitate the comparison of effect sizes. The estimates represent the effect of a change of 1 standard deviation.

We analyzed the risk indicators and their association with mortality with Cox proportional hazard models. We calculated unadjusted hazard ratios and the confidence intervals of each, and subsequently optimized the predefined groups separately, adjusting for age and gender within each group, by means of stacked imputed backward regression, with $P = .2$ as the exclusion threshold.[30] Finally, we combined all remaining variables in a final model using backward regression. We also used least absolute shrinkage and selection operator (LASSO) penalized regression as implemented in R (R Development Core Team, Vienna, Austria) in the package "penalized" to validate the results from backward regressions.[31,32] We evaluated the proportional hazard assumption using Schoenfeld residuals.[33] Variables that did not fulfill the proportional hazards assumption in the imputed datasets were modeled with piecewise constant Heaviside functions.[34] We considered $P < .05$ from 2-sided tests as statistically significant.

We used time-dependent receiver operating characteristic (ROC) curves to compare the predictive performance of the different variable groups over time.[35] ROC curves describe the relationship between sensitivity (true positive rate) and 1 − specificity (false positive rate) for all possible cutoffs of a marker to distinguish between high-risk participants and low-risk participants. We also computed the C-index, the probability that a participant who dies on any given day during a specified time interval has a higher predictive score than one who survives beyond that day. For this part of the analysis, we accounted for residual time dependency using Schoenfeld smoothing.[35] We estimated confidence intervals using cross-validation in 500 bootstrap samples.

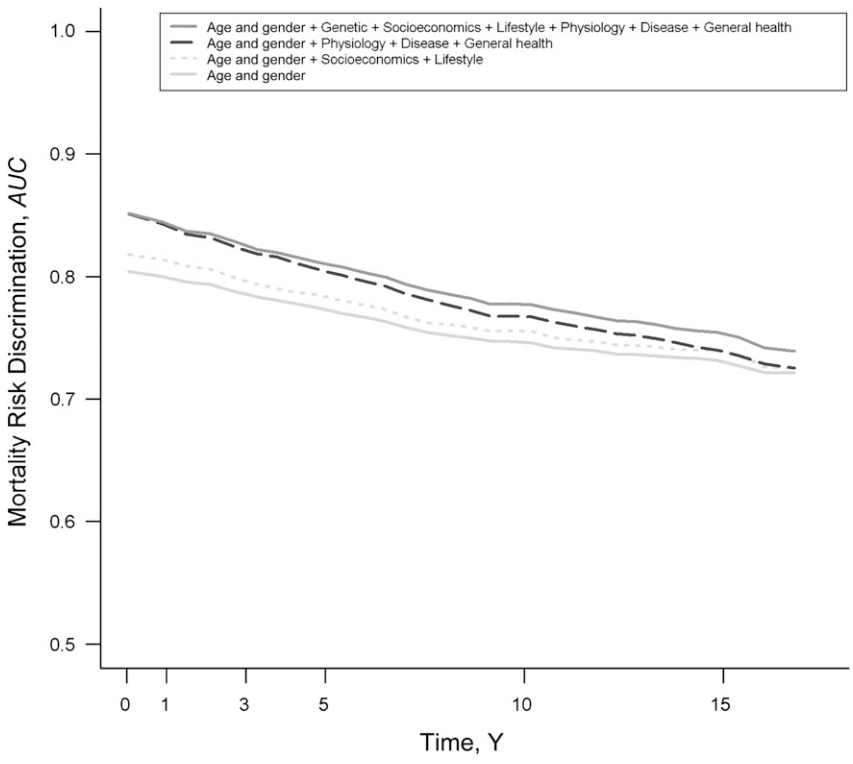
RESULTS

At baseline, participants were on average 69 years of age (range = 55–99 years; Table 1) and 59% were female. Of the 5974 participants, 3174 died (mortality rate = 4.2 per 100 person-years) during the follow-up period. From 162 a priori identified risk factors of mortality (supplemental Table C), backward regression retained 108 variables independently in the final model. Of these, 36 were significantly related to mortality ($P < .05$) as independent risk factors (Table 1). Age (hazard ratio [HR] = 1.09; 95% confidence interval [CI] = 1.08, 1.10) and female gender (HR = 0.71; 95% CI = 0.62, 0.81) were strongly associated with mortality (Table 1).

Of the candidate genes, a priori identified in the literature, *APOE,* insulin-like growth factor 1 receptor (*IGF1R*), and Werner syndrome, RecQ helicase-like (*WRN*) were significant and independent predictors of mortality. From the 50 independent loci, identified from the meta-analysis of 8 discovery genome-wide association studies on time to death, 2 SNPs in the neighborhood of tripartite motif-containing 32 (*TRIM32*) and methionine adenosyltransferase II, β (*MAT2B*) were associated with mortality.

Social class and living in serviced housing were independently associated with risk of death. Smoking status and pack-years as well as energy intake were also associated with mortality. The physiological measures blood pressure, body mass index, and waist circumference, and, in particular, the risk indicators assessed in blood (such as erythrocyte sedimentation rate, leukocytes, creatinine, C-reactive protein, and total cholesterol) or with imaging (such as bone mineral density of the femoral neck and aortic calcification) were all independently related to mortality. Diabetes, cardiac diseases, Parkinson's disease, cancer, and cognitive function remained independently associated with death. Self-perceived comparative health was a good indicator of mortality risk, as were unintentional weight loss and serious illness in the past 5 years.

The predictive power of the variable groups is best explained in 2 ways. First, Figures 1 and 2 show the development of predictive quality



Note. AUC = area under the curve.

FIGURE 2—Time-dependent receiver-operating characteristic curves for prediction of mortality for age and gender and differently combined models: The Rotterdam Study, 1990–2009.

over time. For each point during follow-up, the graphs depict the respective time-dependent area under the curve of a given variable group. Next, Table 2 quantifies the predictive quality for 5 specific prediction intervals (1, 3, 5, 10, and 15 years).

Figure 1 shows that over time, all variable groups except genetic risk markers exhibited decreasing ability to predict death. Figure 1 and Table 2 also demonstrate that prediction based solely on age and gender consistently outperformed all other groups of variables (C-statistic for 15-year follow-up [C_{15y}] = 0.76; 95% CI = 0.75, 0.77; Table 2). Physiological risk and socioeconomic characteristics each predicted mortality equally well over 15 years (each $C_{15y} = 0.72$; 95% CI = 0.71, 0.73). Although of significantly less predictive quality, the C-index of genetic risk markers was still better than chance ($C_{15y} = 0.56$; 95% CI = 0.55, 0.57).

Figure 2 shows the performance of the age- and gender-adjusted model compared with the fully adjusted model ($C_{15y} = 0.80$; 95% CI = 0.79, 0.81). Whereas adding socioeconomic and lifestyle information to age and gender ($C_{15y} = 0.77$; 95% CI = 0.77, 0.78) only slightly improved the predictive quality, the combination of age, gender, general health, disease, and physiology ($C_{15y} = 0.79$; 95% CI = 0.78, 0.80) predicted mortality almost as well as all covariates that remained after backward regression.

To allow comparison with other studies of cause-specific mortality and different population health status, we report the associations of the final model stratified by prevalent, baseline disease status and for cardiovascular disease mortality in Table D (available as a supplement to the online version of this article at http://www.ajph.org).

DISCUSSION

From a set of 162 established risk factors and candidate SNPs for mortality, we identified 36 (31 nongenetic, 5 genetic) independent and significant predictors of mortality. Specific genetic factors were independently associated with mortality and jointly predicted mortality better than chance. However, genetic information added little to age, gender, and other traditional predictors of mortality.

This analysis confirms prior findings that multiple diseases, as well as socioeconomics and lifestyle, jointly influence mortality in the aging adult population.[3,6] Numerous predictors remained independently and significantly associated with mortality. Although several markers of prevalent disease remained associated with mortality, their risk ratios were attenuated. Others, such as prevalent dementia, cerebrovascular accidents, and transient ischemic attacks, were not more effective at predicting mortality than indicators of disease severity and subclinical disease such as the Mini-Mental State Examination and serum C-reactive protein.

Several specific SNPs were independently associated with mortality in our analysis. In accordance with the literature, each additional copy of the *APOE* ε4 allele increased mortality in our study cohort.[1,36,37] Similarly, the *IGF1R* and *WRN* genes have been described before as being associated with longevity and aging, respectively, via improving stress resistance, innate immunity, metabolic maintenance and repair of DNA.[1] This study also confirmed 2 novel loci in the vicinity of the *TRIM32* and *MAT2B* genes as being associated with death. These 2 loci were identified from the pool of SNPs identified by the meta-analysis of the discovery cohorts in the genome-wide association studies. These genes have recently been associated with cancer proliferation.[38,39]

Interestingly, unlike the predictive ability of all other domains, the ability of genetic markers to predict death remained constant during the course of follow-up. The stability of the predictive power of the SNPs observed in this study is probably due to the permanent nature of the genetic makeup. Other variables showed decreasing predictive power over time that can be explained by changes in the value of these variables during the course of follow-up.

Our results support the view that specific SNPs can be identified that are associated with mortality and might be used for risk prediction.[40,41] The results also show, however, that these common SNPs have very limited predictive power and that, especially when used in combination

TABLE 2—C-Index for Different Combinations of Risk Factors at Different Time Points During 15-Year Follow-Up: The Rotterdam Study, 1990–2009

	0–1 Years, C-Index (95% CI)	0–3 Years, C-Index (95% CI)	0–5 Years, C-Index (95% CI)	0–10 Years, C-Index (95% CI)	0–15 Years, C-Index (95% CI)
Age and gender	0.80 (0.78, 0.82)	0.80 (0.78, 0.81)	0.79 (0.77, 0.8)	0.77 (0.76, 0.78)	0.76 (0.75, 0.77)
Genetics	0.55 (0.53, 0.58)	0.55 (0.53, 0.58)	0.55 (0.54, 0.57)	0.56 (0.54, 0.57)	0.56 (0.55, 0.57)
Socioeconomics	0.79 (0.77, 0.81)	0.78 (0.76, 0.8)	0.76 (0.75, 0.78)	0.73 (0.72, 0.74)	0.72 (0.71, 0.73)
Lifestyle	0.64 (0.62, 0.66)	0.63 (0.62, 0.65)	0.62 (0.61, 0.64)	0.60 (0.59, 0.61)	0.59 (0.58, 0.6)
General health	0.79 (0.76, 0.82)	0.77 (0.75, 0.8)	0.75 (0.73, 0.77)	0.71 (0.7, 0.72)	0.68 (0.67, 0.69)
Disease	0.78 (0.74, 0.82)	0.76 (0.73, 0.79)	0.75 (0.72, 0.77)	0.71 (0.7, 0.72)	0.69 (0.68, 0.7)
Physiology	0.79 (0.75, 0.82)	0.77 (0.74, 0.8)	0.76 (0.73, 0.78)	0.73 (0.72, 0.75)	0.72 (0.71, 0.73)
Age and gender + socioeconomics and lifestyle	0.82 (0.79, 0.84)	0.81 (0.79, 0.83)	0.80 (0.79, 0.82)	0.78 (0.78, 0.79)	0.77 (0.77, 0.78)
Age and gender + general health + disease + physiology	0.85 (0.82, 0.87)	0.84 (0.82, 0.86)	0.83 (0.81, 0.84)	0.81 (0.8, 0.82)	0.79 (0.78, 0.8)
Full model: age and gender + socioeconomics + lifestyle + general health + disease + physiology + genetics	0.85 (0.83, 0.88)	0.84 (0.82, 0.86)	0.83 (0.82, 0.85)	0.81 (0.81, 0.82)	0.80 (0.79, 0.81)

Note. CI = confidence interval. The C-indices in this table are the areas under the curve of the graphs in Figures 1 and 2 for 5 different time intervals (1, 3, 5, 10, and 15 y). A C-index of 0.50 indicates a prediction of mortality that is no better than chance, whereas a C-index of 1.0 reflects perfect predictive quality.

with traditional risk factors, they contribute very little, if anything at all, to improve the prediction of death in the general population aged 55 years and older.

Another finding relates to age and gender as predictors of mortality. In our study, the relative risk for mortality per year of age in the univariate model was only reduced by 25% in the fully adjusted model. This is compatible with the idea that aging is not merely the clinical manifestation of disease but an underlying, disease-independent accumulation of pathophysiological changes that favor mortality over time.[42–44]

The gender differences in this study were not due to differences in prevalent diseases at the onset of the study. Females exhibited even stronger reduced risks after adjustment for other risk factors. This strongly suggests that the gender difference in survival cannot be explained by differences in health behavior and disease at baseline.[45] We can only speculate that the survival benefits of females can be found in gender effects or different genetic origins not accounted for in this analysis. One of the potential genetic candidates that could contribute to the female survival advantage is the X chromosome.[46] Another genetic candidate is mitochondrial DNA (mtDNA). It has been suggested that, because mtDNA is inherited from the maternal line, a possible intergenomic conflict between mtDNA and nuclear DNA favors female survival.[47] We could not include markers on the X chromosome and mtDNA because the X chromosome is not commonly analyzed in all the discovery cohorts and mtDNA is also not available on all genotyping platforms.

Prediction of mortality by age and gender improved by only 5% upon inclusion of all independent mortality predictors over the entire 15 years of follow-up. This is particularly interesting considering the multitude of independent risk factors identified in this study, and because all groups contributed significant variables. From the time-dependent area-under-the-curve analyses, 2 observations are particularly noteworthy: first, most of the additional information beyond age and gender stemmed from indicators of physiology, disease, and general health; second, although socioeconomic factors were equally good at predicting mortality as indicators of disease, the combination of socioeconomics, lifestyle, and genetic markers contributed visibly to the explanation of mortality risk only after 10 years of follow-up and beyond. Thus, although socioeconomics and lifestyle were associated to mortality, they seemingly exerted their effects on mortality through physiological risk indicators and disease rather than by acting independently on mortality.[48] This underscores the importance of socioeconomics and living conditions for public policy aimed at reducing health inequalities.

To summarize the findings reported here and in other studies, one can insinuate a cascade from gene to individual health to death, in which every step is accompanied by environmental influences, some of which are controlled by the individual (such as physical activity and obesity) whereas others are defined by the individual's living circumstances, cultural heritage and surroundings. Figure A (available as a supplement to the online version of this article at http://www.ajph.org) illustrates at which stage during the course of aging different interventions (e.g., improvements in living circumstances or the introduction of a new therapeutic drug) can feasibly act and which health gains could be expected.

Limitations

Caution is needed in interpreting this study. It was not our aim to evaluate the size of the mortality risk associated with single risk factors. Several of the markers in this analysis describe the same underlying construct (e.g., body composition). Therefore, the specific relative risks must be interpreted cautiously. Other important aspects of health such as physical activity and mental health are barely represented among the risk factors analyzed, as only "riding a bike" and "self-perceived comparative

health" were available to approximate these important dimensions of health. Another limitation concerns the genetic markers used in this study. We included only autosomal SNPs. Genetic risk is transferred through several other mechanisms, including DNA methylation, copy number variations, and mitochondrial DNA. Furthermore, because this study has not been replicated externally, it probably cannot be used for constructing a risk score. We did not seek external replication because of the multitude of specific risk factors and instead relied on bootstrapping and cross-validation for guiding the LASSO analysis and the estimation of the C-index. At the same time, the selection of SNPs is among the strengths of this study, as the SNPs were identified from 2 sources, independent of the population under study. Other strengths are related to the multitude of risk factors and the prospective design with long follow-up.

Conclusions

We found 36 variables that independently and significantly predicted mortality in the Rotterdam Study population. Adding further risk indicators to age and gender improved our ability to predict death, but the gain in predictive quality was modest, particularly in the long run. Surprisingly, specific genetic risk factors, independently and as a group, predicted mortality, but their added value to conventional predictors of mortality was low. Our findings also support the importance of primary prevention in the areas of socioeconomics and lifestyle, as we could illustrate how these risk factors continuously influence mortality risk independently and through their impact on physiological risk status and disease. ■

About the Authors

Stefan Walter, Johan Mackenbach, and Stefan Lhachimi are with the Department of Public Health; M. Arfan Ikram, Albert Hofman, and Henning Tiemeier are with the Department of Epidemiology; and André G. Uitterlinden is with the Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. Zoltán Vokó is with the Department of Health Policy and Health Economics, Institute of Economics, Faculty of Social Sciences, Eötvös Loránd University, Budapest, Hungary. Anne B. Newman is with the Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA. Joanne M. Murabito is with the National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA. Melissa E. Garcia and Lenore J. Launer are with the Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, National Institutes of Health (NIH), Bethesda, MD. Vilmundur Gudnason is with the Icelandic Heart Association, University of Iceland, Reykjavik, Iceland. Toshiko Tanaka is with the Clinical Research Branch, National Institute on Aging, Baltimore, MD. Gregory J. Tranah is with the California Pacific Medical Center, San Francisco, CA. Henri Wallaschofski is with the Institute of Clinical Chemistry and Laboratory Medicine and Thomas Kocher is with the Dental School, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany. Nora Franceschini is with the Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill. Maarten Schipper is with the Rijksinstituut voor Volksgezondheid en Milieu (RIVM), Bilthoven, The Netherlands.

Correspondence should be sent to Henning Tiemeier, MD, PhD, Erasmus Medical Center, Department of Epidemiology, PO Box 2040, 3000 CA Rotterdam, The Netherlands (e-mail: h.tiemeier@erasmusmc.nl). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This article was accepted November 21, 2011.

Contributors

S. Walter, Z. Vokó, M. Schipper, and H. Tiemeier conceptualized the study and led the design. S. Walter, J. Mackenbach, A. Hofman, and H. Tiemeier led analysis and interpretation of data and drafted the article. All authors critically revised the article for important intellectual content and approved the final draft.

Acknowledgments

The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam, The Netherlands Organization for Scientific Research, The Netherlands Organization for Health Research and Development, the Research Institute for Diseases in the Elderly, the Ministry of Education, Culture and Science, the Ministry of Health, Welfare and Sports, and the European Commission and the Municipality of Rotterdam. This research was supported in part by the Intramural Research Program, NIH, National Institute on Aging. Stefan Walter is financed by Netspar and the Netspar theme "Living Longer in Good Health." The work of Henning Tiemeier is supported by Vidi (grant 017.106.370).

Human Participant Protection

The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus Medical Center. All participants provided written informed consent.

References

1. Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet.* 2006;7(6):436–448.

2. Oeppen J, Vaupel Demography JW. Broken limits to life expectancy. *Science.* 2002;296(5570):1029–1031.

3. Fried LP, Kronmal RA, Newman AB, et al. Risk factors for 5-year mortality in older adults: The Cardiovascular Health Study. *JAMA.* 1998;279(8):585–592.

4. Noale M, Minicuci N, Bardage C, et al. Predictors of mortality: an international comparison of socio-demographic and health characteristics from six longitudinal studies on aging: the CLESA project. *Exp Gerontol.* 2005;40(1–2):89–99.

5. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA.* 2006;295(7):801–808.

6. Newman AB, Sachs MC, Arnold AM, et al. Total and cause-specific mortality in the cardiovascular health study. *J Gerontol A Biol Sci Med Sci.* 2009;64(12):1251–1261.

7. Baer HJ, Glynn RJ, Hu FB, et al. Risk factors for mortality in the nurses' health study: a competing risks analysis. *Am J Epidemiol.* 2011;173(3):319–329.

8. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol.* 1991;7(4):403–422.

9. Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol.* 2009;24(9):553–572.

10. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J. Health Soc Behav.* 1995;(spec no.): 80–94.

11. Psaty BM, O'Donnell CJ, Gudnason V, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet.* 2009; 2(1):73–80.

12. Slooter AJC, Cruts M, Kalmijn S, et al. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: The Rotterdam Study. *Arch Neurol.* 1998;55(7):964–968.

13. Klipstein-Grobusch K, den Breeijen JH, Goldbohm RA, et al. Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. *Eur J Clin Nutr.* 1998;52(8):588–596.

14. *Nevo Food Composition Table From the Netherlands Nutrient Databank.* Den Haag, The Netherlands: Voorlichtingsbureau voor de Voeding; 1993.

15. Visscher TL, Seidell JC, Molarius A, van der Kuip D, Hofman A, Witteman JC. A comparison of body mass index, waist-hip ratio and waist circumference as predictors of all-cause mortality among the elderly: The Rotterdam Study. *Int J Obes Relat Metab Disord.* 2001; 25(11):1730–1735.

16. Hak AE, Pols HA, Visser TJ, Drexhage HA, Hofman A, Witteman JC. Subclinical hypothyroidism is an independent risk factor for atherosclerosis and myocardial infarction in elderly women: The Rotterdam Study. *Ann Intern Med.* 2000;132(4):270–278.

17. Ikram MK, de Jong FJ, Vingerling JR, et al. Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? The Rotterdam Study. *Invest Ophthalmol Vis Sci.* 2004;45(7):2129–2134.

18. Hak AE, Pols HAP, Stehouwer CDA, et al. Markers of inflammation and cellular adhesion molecules in relation to insulin resistance in nondiabetic elderly: The Rotterdam Study. *J Clin Endocrinol Metab.* 2001;86(9): 4398–4405.

19. Bos MJ, Koudstaal PJ, Hofman A, Witteman JCM, Breteler MMB. Uric acid is a risk factor for myocardial infarction and stroke: The Rotterdam Study. *Stroke.* 2006;37(6):1503–1507.

20. Burger H, van Daele PL, Algra D, et al. The association between age and bone mineral density in men and women aged 55 years and over: The Rotterdam Study. *Bone Miner.* 1994;25(1):1–13.

21. Ikram MK, Witteman JCM, Vingerling JR, Breteler MMB, Hofman A, de Jong PTVM. Retinal vessel diameters and risk of hypertension: The Rotterdam Study. *Hypertension.* 2006;47(2):189–194.

22. Ott A, Stolk RP, Hofman A, van Harskamp F, Grobbee DE, Breteler MM. Association of diabetes mellitus and dementia: The Rotterdam Study. *Diabetologia.* 1996;39(11):1392–1397.

23. Mosterd A, Hoes AW, de Bruyne MC, et al. Prevalence of heart failure and left ventricular dysfunction in the general population: The Rotterdam Study. *Eur Heart J.* 1999;20(6):447–455.

24. Hofman A, van Duijn CM, Franco OH, et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol.* 2011;26:657–686.

25. Bots ML, Looman SJ, Koudstaal PJ, Hofman A, Hoes AW, Grobbee DE. Prevalence of stroke in the general population: The Rotterdam Study. *Stroke.* 1996;27(9):1499–1501.

26. van der Klift M, De Laet CEDH, McCloskey EV, Hofman A, Pols HAP. The incidence of vertebral fractures in men and women: The Rotterdam Study. *J Bone Miner Res.* 2002;17(6):1051–1056.

27. van der Klift M, de Laet CE, Coebergh JW, Hofman A, Pols HA. Bone mineral density and the risk of breast cancer: The Rotterdam Study. *Bone.* 2003;32(3):211–216.

28. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. *J Rheumatol.* 1982;9(5):789–793.

29. Lawton MP, Moss M, Fulcomer M, Kleban MH. A research and service oriented multilevel assessment instrument. *J Gerontol.* 1982;37(1):91–99.

30. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *Stat Med.* 2008;27(17):3227–3246.

31. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* 1997;16(4):385–395.

32. Goeman J. L1 penalized estimation in the Cox proportional hazards model. *Biome J.* 2010;52(1)70–84.

33. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika.* 1982;69(1):239–241.

34. Kleinbaum DG, Klein M. *Survival Analysis: A Self-Learning Text.* Vol 2. New York, NY: Springer; 2005.

35. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics.* 2005;61(1):92–105.

36. Ewbank DC. Differences in the association between apolipoprotein E genotype and mortality across populations. *J Gerontol A Biol Sci Med Sci.* 2007;62(8):899–907.

37. Vaupel JW. Biodemography of human ageing. *Nature.* 2010;464(7288):536–542.

38. Kano S, Miyajima N, Fukuda S, Hatakeyama S. Tripartite motif protein 32 facilitates cell growth and migration via degradation of Abl-interactor 2. *Cancer Res.* 2008;68(14):5572–5580.

39. Lu SC, Mato JM. S-Adenosylmethionine in cell growth, apoptosis and liver cancer. *J Gastroenterol Hepatol.* 2008;23(suppl 1):S73–S77.

40. Lunetta KL, D'Agostino RB Sr, Karasik D, et al. Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. *BMC Med Genet.* 2007;8(suppl 1):S13.

41. Kraft P, Hunter DJ. Genetic risk prediction—are we there yet? *N Engl J Med.* 2009;360(17):1701–1703.

42. Franco OH, Karnik K, Osborne G, Ordovas JM, Catt M, van der Ouderaa F. Changing course in ageing research: The Healthy Ageing Phenotype. *Maturitas.* 2009;63(1):13–19.

43. Franceschi C, Capri M, Monti D, et al. Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans. *Mech Ageing Dev.* 2007;128(1):92–105.

44. Hayflick L. The future of ageing. *Nature.* 2000;408(6809):267–269.

45. Liang J, Bennett JM, Sugisawa H, Kobayashi E, Fukaya T. Gender differences in old age mortality: roles of health behavior and baseline health status. *J Clin Epidemiol.* 2003;56(6):572–582.

46. Christensen K, Orstavik KH, Vaupel JW. The X chromosome and the female survival advantage: an example of the intersection between genetics, epidemiology and demography. *Ann N Y Acad Sci.* 2001;954:175–183.

47. Rand DM. Mitochondrial genetics of aging: intergenomic conflict resolution. *Sci Aging Knowledge Environ.* 2005;2005(45):re5.

48. Steenland K, Henley J, Calle E, Thun M. Individual- and area-level socioeconomic status variables as predictors of mortality in a cohort of 179,383 persons. *Am J Epidemiol.* 2004;159(11):1047–1056.

Atherosclerosis 222 (2012) 138–147

Contents lists available at SciVerse ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

Genetic determinants of the ankle-brachial index: A meta-analysis of a cardiovascular candidate gene 50K SNP panel in the candidate gene association resource (CARe) consortium

Christina L. Wassel[a,1], Claudia Lamina[b,1], Vijay Nambi[c,1], Stefan Coassin[b,1], Kenneth J. Mukamal[d,1], Santhi K. Ganesh[e], David R. Jacobs Jr.[f], Nora Franceschini[g], George J. Papanicolaou[h], Quince Gibson[i], Lisa R. Yanek[j], Pim van der Harst[k], Jane F. Ferguson[l], Dana C. Crawford[m], Lindsay L. Waite[n], Matthew A. Allison[a], Michael H. Criqui[a], Mary M. McDermott[o], Reena Mehra[p], L. Adrienne Cupples[q], Shih-Jen Hwang[r], Susan Redline[s], Robert C. Kaplan[t], Gerardo Heiss[g], Jerome I. Rotter[u], Eric Boerwinkle[v], Herman A. Taylor[w], Luis H. Eraso[x], Margot Haun[b], Mingyao Li[y], Christa Meisinger[z], Jeffrey R. O'Connell[i], Alan R. Shuldiner[i,A], Anne Tybjærg-Hansen[B], Ruth Frikke-Schmidt[B], Barbara Kollerits[b], Barbara Rantner[b,C], Benjamin Dieplinger[D], Marietta Stadler[E], Thomas Mueller[D], Meinhard Haltmayer[D], Peter Klein-Weigel[F], Monika Summerer[b], H.-Erich Wichmann[G,H,I], Folkert W. Asselbergs[J,K,L], Gerjan Navis[M], Irene Mateo Leach[k], Kristin Brown-Gentry[N], Robert Goodloe[N], Themistocles L. Assimes[O], Diane M. Becker[j], John P. Cooke[O], Devin M. Absher[n], Jeffrey W. Olin[P], Braxton D. Mitchell[i], Muredach P. Reilly[l], Emile R. Mohler III[Q,1], Kari E. North[R,1], Alexander P. Reiner[S,1], Florian Kronenberg[b,1], Joanne M. Murabito[T,1,*]

[a] *University of California San Diego, Department of Family and Preventive Medicine, Division of Preventive Medicine, La Jolla, CA, USA*
[b] *Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria*
[c] *Baylor College of Medicine, Department of Medicine, Section in Cardiology, Houston, TX, USA*
[d] *Division of General Medicine & Primary Care, Beth Israel Deaconess Medical Center, Boston, MA, USA*
[e] *Division of Cardiovascular Medicine, Deparment of Internal Medicine, Univeristy of Michigan Health Care System, Ann Arbor, MI, USA*
[f] *Department of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA*
[g] *University of North Carolina, Department of Epidemiology, Gillings Global School of Public Health, Chapel Hill, NC, USA*
[h] *Division of Cardiovascular Sciences, National Heart, Lung and Blood Institute, Bethesda, MD, USA*
[i] *Division of Endocrinology, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA*
[j] *Johns Hopkins University School of Medicine, Division of General Internal Medicine, Baltimore, MD, USA*
[k] *Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands*
[l] *Perelman School of Medicine at the University of Pennsylvania, Cardiovascular Institute, Philadelphia, PA, USA*
[m] *Department of Molecular Physiology and Biophysics, The Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA*
[n] *HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA*
[o] *Northwestern University Feinberg School of Medicine, Department of Medicine, Chicago, IL, USA*
[p] *Case Western Reserve University, University Hospitals Case Medical Center, Cleveland, OH, USA*
[q] *Boston University School of Public Health, Department of Biostatistics, Boston, MA, USA*
[r] *The Framingham Heart Study, Center for Population Study, National Heart Lung and Blood Institute, Framingham, MA, USA*
[s] *Brigham and Women's Hospital, Division of Sleep Medicine,Harvard Medical School, Boston, MA, USA*
[t] *Albert Einstein College of Medicine, Department of Epidemiology and Population Health, Bronx, NY, USA*
[u] *Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA*
[v] *University of Texas Health Science Center at Houston, Department of Epidemiology, Human Genetics and Environmental Sciences, Houston, TX, USA*
[w] *University of Mississippi Medical Center AND Jackson State University, Jackson, MS, USA*
[x] *Jefferson Vascular Center, Thomas Jefferson University Hospital, Philadelphia, PA, USA*
[y] *Perelman School of Medicine at the University of Pennsylvania, Center for Clinical Epidemiology and Biostatistics, Philadelphia, PA, USA*
[z] *Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany*
[A] *Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, MD, USA*
[B] *Department Clin. Biochem., Rigshospitalet, Copenhagen University Hospitals, Copenhagen, Denmark*
[C] *Department of Vascular Surgery, Innsbruck Medical University, Innsbruck, Austria*

* Corresponding author at: 73 Mt. Wayte Avenue, Framingham, MA 01702, USA. Tel.: +1 508 935 3461; fax: +1 508 626 1262.
E-mail address: Murabito@bu.edu (J.M. Murabito).
[1] Equal work was done by these authors for this paper.

doi:10.1016/j.atherosclerosis.2012.01.039

[D] Department of Laboratory Medicine, Konventhospital Barmherzige Brueder Linz, Linz, Austria
[E] Hietzing Hospital, 3rd Medical Department of Metabolic Diseases and Nephrology, Vienna, Austria
[F] Klinik für Innere Medizin – Schwerpunkt Angiologie, DRK-Kliniken Berlin Mitte, Berlin, Germany
[G] Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[H] Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany
[I] Klinikum Grosshadern, Munich, Germany
[J] Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands
[K] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[L] Department of Medical Genetics, Biomedical Genetics, University Medical Center, Utrecht, The Netherlands
[M] Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[N] The Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA
[O] Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA
[P] Zena and Michael A. Wiener CardiovascularInstitute & Marie-Joseñe and Henry R. Kravis Center for Cardiovascular Health, Mount Sinai School of Medicine, New York, NY, USA
[Q] Perelman School of Medicine at the University of Pennsylvania, Cardiovascular Division, Vascular Medicine Section, Philadelphia, PA, USA
[R] Department of Epidemiology and Carolina Center for Genome Sciences, The University of North Carolina, Chapel Hill, NC, USA
[S] Department of Epidemiology, University of Washington School of Public Health, Seattle, WA, USA
[T] Boston University School of Medicine, Department of Medicine, Section of General Internal Medicine; NHLBI's Framingham Heart Study, Boston, MA, USA

ARTICLE INFO

Article history:
Received 6 October 2011
Received in revised form 23 January 2012
Accepted 26 January 2012
Available online 2 February 2012

Keywords:
Ankle brachial index
Peripheral artery disease
Genetics
Candidate gene array
Meta-analysis
Ethnicity

ABSTRACT

Background: Candidate gene association studies for peripheral artery disease (PAD), including subclinical disease assessed with the ankle-brachial index (ABI), have been limited by the modest number of genes examined. We conducted a two stage meta-analysis of ~50,000 SNPs across ~2100 candidate genes to identify genetic variants for ABI.

Methods and results: We studied subjects of European ancestry from 8 studies (n = 21,547, 55% women, mean age 44–73 years) and African American ancestry from 5 studies (n = 7267, 60% women, mean age 41–73 years) involved in the candidate gene association resource (CARe) consortium. In each ethnic group, additive genetic models were used (with each additional copy of the minor allele corresponding to the given beta) to test each SNP for association with continuous ABI (excluding ABI > 1.40) and PAD (defined as ABI < 0.90) using linear or logistic regression with adjustment for known PAD risk factors and population stratification. We then conducted a fixed-effects inverse-variance weighted meta-analyses considering a $p < 2 \times 10^{-6}$ to denote statistical significance.

Results: In the European ancestry discovery meta-analyses, rs2171209 in *SYTL3* ($\beta = -0.007$, $p = 6.02 \times 10^{-7}$) and rs290481 in *TCF7L2* ($\beta = -0.008$, $p = 7.01 \times 10^{-7}$) were significantly associated with ABI. None of the SNP associations for PAD were significant, though a SNP in *CYP2B6* ($p = 4.99 \times 10^{-5}$) was among the strongest associations. These 3 genes are linked to key PAD risk factors (lipoprotein(a), type 2 diabetes, and smoking behavior, respectively). We sought replication in 6 population-based and 3 clinical samples (n = 15,440) for rs290481 and rs2171209. However, in the replication stage (rs2171209, $p = 0.75$; rs290481, $p = 0.19$) and in the combined discovery and replication analysis the SNP–ABI associations were no longer significant (rs2171209, $p = 1.14 \times 10^{-3}$; rs290481, $p = 8.88 \times 10^{-5}$). In African Americans, none of the SNP associations for ABI or PAD achieved an experiment-wide level of significance.

Conclusions: Genetic determinants of ABI and PAD remain elusive. Follow-up of these preliminary findings may uncover important biology given the known gene-risk factor associations. New and more powerful approaches to PAD gene discovery are warranted.

1. Introduction

Peripheral artery disease (PAD) is associated with an increased risk for incident cardiovascular disease events and mortality [1,2]. In the reduction of atherothrombosis for continued health (REACH) registry, almost two-thirds of the individuals with PAD had concomitant clinically evident atherosclerotic disease in the cerebrovascular or coronary artery disease (CAD) territories whereas only one-quarter of the individuals with coronary disease had clinically evident atherosclerotic involvement of other arterial beds [3]. While PAD and CAD share many common risk factors, cigarette smoking and type 2 diabetes are stronger risk factors for PAD than for coronary artery disease [4]. The variable distribution of the burden of atherosclerosis across vascular beds among subjects at risk suggests that other factors exist, including possibly genetic factors, that may contribute to the predilection of atherosclerosis to develop in a given anatomic location. Currently, little is known about the genetic susceptibility to PAD but familial aggregation of PAD and heritability estimates suggest a significant genetic contribution [5–10].

The ankle brachial index (ABI) is an easy and reliable diagnostic test used to detect symptomatic as well as asymptomatic PAD [11]. A genome-wide linkage scan for ABI identified several potential candidate genes under six linkage signals in pathways of inflammation, coagulation, blood pressure regulation, and lipid metabolism [8]. A recent large genome-wide association study (GWAS) meta-analysis of European descent participants found variants in the 9p21 locus significantly associated with ABI [12]. However, there have been few candidate gene association studies for PAD, most of which have been limited by small sample size, modest number of genes examined, and lack of robust independent replication of initial findings [13]. We conducted a two stage large scale candidate gene association study of the ~2100 candidate genes included in a cardiovascular gene-centric 50K single nucleotide polymorphism (SNP) array [14] within the candidate gene association resource (CARe) consortium that included 21,547 individuals of European ancestry (1190 with ABI < 0.9) from eight cohorts and 7267 African American individuals (594 with ABI < 0.90) from six cohorts. In the second stage of the investigation, we sought to replicate our significant associations among individuals of European ancestry in 13,524 individuals from population-based cohort studies and 1916 individuals from clinically based studies. We hypothesized that this approach would lead to the identification of novel genetic variants associated with ABI and PAD (as defined by an ABI < 0.90). Furthermore, we hypothesized that some variants may influence both PAD risk factors and PAD itself, as has been observed

in genome-wide studies of lipids and coronary artery disease [15,16].

2. Methods

2.1. Discovery studies: CARe consortium and additional studies

The CARe consortium (http://public.nhlbi.nih.gov/GeneticsGenomics/home/care.aspx) was funded by the National Heart Lung and Blood Institute (NHLBI) in 2006 to explore the association of a custom cardiovascular gene centric SNP array [14] with a broad set of cardiovascular, metabolic, and inflammatory phenotypes collected across nine longitudinal cohort studies [17]. The following CARe consortium studies contributed data to the present analysis (Tables 1A and 1B): atherosclerosis risk in communities study (ARIC, $n = 9031$ European Americans, $n = 2853$ African Americans), the Cleveland family study (CFS, $n = 275$ European Americans, $n = 365$ African Americans), the cardiovascular health study (CHS, $n = 3826$ European Americans, $n = 722$ African Americans), the Framingham heart study (FHS, $n = 2701$ European Americans), the Jackson heart study (JHS, $n = 1734$ African Americans), and the multi-ethnic study of atherosclerosis (MESA, $n = 2280$ European Americans, $n = 1593$ African Americans). The Amish study ($n = 1008$), the cooperative research in the region of Augsburg (KORA F3, $n = 1807$), and the Penn diabetes heart study (PDHS, $n = 622$) cohort all used the same cardiovascular gene centric SNP array and joined the discovery stage analyses conducted in European Americans.

Description of each study is provided in the Supplementary Materials. For all studies, each participant self-identified as either White (European, European American) or African American and provided written informed consent. The Institutional Review Board at the parent institution for each respective study approved the study protocols.

The characteristics of the discovery study samples at the time of ABI measurement are presented in Tables 1A and 1B by ethnic group. More than half were women and the mean age ranged from 44 years (CFS) to 73 years (CHS) in samples of European ancestry, and from 41 years (CFS) to 73 years (CHS) in African Americans. The mean ABI was 1.10 and the prevalence of PAD (ABI < 0.90) varied across studies, ranging from 4% to 12% in European Americans and from 5% to 21% in African Americans. Risk factor burden appeared greater in African Americans, as demonstrated by higher prevalence of hypertension, type 2 diabetes, and obesity compared to European ancestry participants.

2.2. ABI phenotypes

The details of the ABI protocol for each study are provided in Supplementary Materials, Supplementary Methods Table 1. For each leg, the systolic blood pressure at each ankle was divided by the systolic blood pressure in the arm. In the ARIC study, ABI was measured in only one leg and one arm chosen at random. The lower of the two ABIs calculated with each leg was used for analyses with the exception of the Amish study which used the average of an individual's two ABIs in the analyses.

We defined two PAD phenotypes for genetic association analyses. First, we used the continuous range of ABI ≤ 1.40. Next, we defined PAD as ABI < 0.90 and conducted a case (ABI < 0.9)/control comparison (ABI ≥ 0.90 and < 1.40) analysis. Participants with an ABI > 1.40 were excluded as these subjects likely had medial artery calcification and therefore the artery would not be compressible to allow for determination of pressure in the artery. Excluding participants with ABI did not truncate the distribution substantially, and ABI was still normally distributed.

Table 1A
Characteristics of discovery samples at the time of ankle brachial index (ABI) measurement. European and European Americans.[a]

Characteristic mean (SD) or n (%)	ARIC $n = 9031$	CFS $n = 275$	CHS $n = 3826$	FHS[b] $n = 2701$	MESA $n = 2280$	Amish study $n = 1008$	KORA F3 $n = 1807$	PDHS $n = 622$
Age (years)	54 ± 6	44 ± 19	73 ± 6	61 ± 9	63 ± 10	44 ± 14	62 ± 11	59 ± 9
Women (%)	4871 (54%)	145 (53%)	2152 (56%)	1453 (54%)	1196 (52%)	484 (48%)	965 (53%)	429 (69%)
ABI (mean)	1.13 ± 0.13	1.08 ± 0.10	1.06 ± 0.16	1.13 ± 0.12	1.11 ± 0.12	1.10 ± 0.10	1.11 ± 0.13	1.12 ± 0.14
PAD (ABI < 0.9) (%)	334 (4%)	12 (4%)	453 (12%)	100 (4%)	105 (5%)	36 (3.6%)	92 (5%)	40 (6.5%)
Hypertension (%)	2410 (27%)	85 (31%)	2136 (56%)	629 (23%)	882 (39%)	125 (12%)	1046 (58%)	361 (58%)
Diabetes (%)	785 (9%)	34 (12%)	547 (14%)	288 (11%)	141 (6%)	6 (0.6%)	168 (9%)	622 (100%)
Body mass index (kg/m^2)	27 ± 5	32 ± 9	26 ± 4	28 ± 5	28 ± 5	27 ± 5	28 ± 5	33 ± 6
Ever smoker (%)	5387 (60%)	120 (44%)	2076 (54%)	1595 (59%)	1274 (56%)	239 (24%)	929 (51%)	361 (58%)
Total cholesterol (mg/dL)	215 ± 41	192 ± 42	211 ± 39	200 ± 36	196 ± 36	209 ± 47	222 ± 40	177 ± 38
LDL cholesterol (mg/dL)	137 ± 38	102 ± 30	130 ± 35	119 ± 33	117 ± 30	139 ± 4	130 ± 33	99 ± 30
HDL cholesterol (mg/dL)	51 ± 17	44 ± 12	54 ± 16	54 ± 17	52 ± 16	56 ± 15	59 ± 17	46 ± 13
Triglyceride (mg/dL)	137 ± 93	139 ± 101	144 ± 79	137 ± 89	134 ± 92	69 ± 41	171 ± 133	155 ± 88
Lipid lowering meds (%)	297 (3%)	40 (15%)	73 (2%)	559 (21%)	414 (18%)	32 (2%)	247 (14%)	386 (62%)
Claudication (%)	70 (0.8%)	3 (1%)	79 (2%)	45 (2%)	9 (0.4%)	N/A[c]	67 (3.7%)	0 (0%)
Prevalent CVD (%)	0 (0%)	14 (5%)	879 (23%)	110 (4%)	0 (0%)	28 (2.9%)	227 (13%)	0 (0%)

[a] ARIC = atherosclerosis risk in communities, CFS = Cleveland family study, CHS = cardiovascular health study, FHS = Framingham heart study, MESA = multi-ethnic study of atherosclerosis, KORA F3 = cooperative research in the region of Augsburg, PDHS = Penn diabetes heart study.
[b] Includes both original cohort and offspring cohort.
[c] N/A = not measured or assessed in the study.

Table 1B
Characteristics of discovery samples at the time of ankle brachial index (ABI) measurement. African Americans.[a]

Characteristic mean (SD) or n (%)	ARIC $n = 2853$	CFS $n = 365$	CHS $n = 722$	JHS $n = 1734$	MESA $n = 1593$
Age (years)	53 ± 6	41 ± 19	73 ± 6	50 ± 12	62 ± 10
Women (%)	1798 (63%)	208 (57%)	454 (63%)	1051 (61%)	867 (54%)
ABI (mean)	1.1 ± 0.1	1.0 ± 0.1	1.0 ± 0.2	1.1 ± 0.1	1.1 ± 0.1
PAD (ABI < 0.9) (%)	148 (5%)	45 (12%)	153 (21%)	105 (6%)	143 (9%)
Hypertension (%)	1578 (55%)	155 (42%)	520 (72%)	970 (56%)	959 (60%)
Diabetes (%)	543 (19%)	72 (20%)	170 (24%)	251 (14%)	281 (18%)
Body mass index (kg/m^2)	30 ± 6	33 ± 9	28 ± 6	32 ± 7	30 ± 6
Ever smoker (%)	1500 (53%)	175 (48%)	367 (51%)	544 (31%)	860 (54%)
Total cholesterol (mg/dL)	215 ± 45	183 ± 43	209 ± 39	197 ± 39	189 ± 36
LDL cholesterol (mg/dL)	138 ± 43	97 ± 34	128 ± 36	126 ± 36	116 ± 33
HDL cholesterol (mg/dL)	55 ± 18	44 ± 13	58 ± 16	51 ± 14	52 ± 15
Triglyceride (mg/dL)	113 ± 74	103 ± 63	116 ± 63	106 ± 89	105 ± 71
Lipid lowering meds (%)	39 (1%)	50 (14%)	32 (4%)	154 (9%)	252 (16%)
Claudication (%)	18 (0.6%)	3 (0.8%)	11 (2%)	534 (30%)	10 (0.6%)
Prevalent CVD (%)	0 (0%)	11 (3%)	191 (26%)	98 (6%)	0 (0%)

[a] ARIC = atherosclerosis risk in communities, CFS = Cleveland family study, CHS = cardiovascular heart study, JHS = Jackson heart study, MESA = multi-ethnic study of atherosclerosis.

2.3. Genotyping

Genotyping in the CARe cohorts and PDHS was conducted at the Broad Institute using the ITMAT-Broad-CARe (IBC) Illumina iSELECT custom array [14]. The Old Order Amish Study genotyping was also performed using the ITMAT IBC Illumina iSELECT custom array. For KORA F3 (Discovery stage) genotyping with the same array was performed in the Genome Analysis Centre, Helmholtz Zentrum München.

The IBC array was designed to capture genetic variation in loci known or postulated to be associated with cardiovascular disease, metabolic disease and inflammatory diseases [14]. Specifically, a cosmopolitan tagging approach was used to capture genetic diversity across ~2100 candidate genes [14]. Loci were primarily chosen in three groups as follows: (1) 435 loci were chosen areas with a high probability of functional significance, (2) 1349 loci were chosen as having involvement in phenotypes of interest (i.e. cardiovascular disease or cardiovascular disease-related traits such as inflammation, hemostasis, obesity, diabetes) or were well-established loci requiring a number of SNPs for coverage, and (3) 232 lower priority loci were chosen, which also included larger genes [14]. Further details on the IBC array can be found elsewhere [14]. Details of the genotyping and quality control procedures are provided in Supplementary Methods Table 2.

2.4. Statistical analysis

For each study, residuals of ABI stratified by gender and race were created from linear regression models and used as phenotypes in the association analysis; results of gender were pooled but all analyses were stratified by race. The ABI residuals were adjusted for age, clinic site for multi-site studies, principal components (participants of European ancestry) or global European ancestry (African American participants), ever smoking, type 2 diabetes, hypertension (>140/90 or use of anti-hypertensive medication), LDL cholesterol, HDL cholesterol, and body mass index (BMI). In each ethnic group, SNPs were modeled additively, and the association of each SNP with ABI was tested using linear regression. The PDHS study did not use diabetes as a covariate as all subjects were diagnosed with type 2 diabetes. For CFS and FHS, linear mixed effects (LME) models were used to account for familial correlations. Multivariable logistic regression was used to test for the association of each SNP with PAD. For CFS and FHS, generalized estimating equations (GEE) were used to account for familial correlations. The covariate adjustment for PAD was the same as used for the ABI phenotype.

A fixed effects meta-analysis with inverse-variance weighting was then conducted in PLINK V 1.0.6 [18] and Stata V 9.0 (College Station, TX) to combine the results for all studies. The association of each additional copy of the minor allele with ABI was quantified by the regression slope (β), its standard error [SE(β)] and the corresponding p-value. We calculated a meta-analysis odds ratio (OR) for each of the most significant SNP associations for PAD. The meta-analysis OR represents the increase/decrease in odds of PAD for each additional copy of the minor allele of the SNP. We also tested for heterogeneity of study-specific regression parameters using the Cochran's Q statistic in Stata V9.0, and report the p-values for heterogeneity. Associations were considered to be significant on an experiment-wide level at a p-value $\leq 2 \times 10^{-6}$ which was determined based on the estimate of the number of independent tests [19]. SNPs with MAF < 0.01 were excluded.

A gene-based test of association using the meta-analyzed p-values at the discovery stage was performed using the program Versatile Gene-Based Association Study (VEGAS) [20]; http://gump.qimr.edu.au/VEGAS/). The SNPs are matched to genes using the UCSC Genome Browser hg18 assembly with the gene region defined by ±50 kb up-and downstream of the gene. The test is based on the sum of chi-square-statistics and the linkage disequilibrium (LD) of these SNPs is taken into account according to the correlation structure in the HapMap CEU samples. An empirical p-value is provided based on all SNPs, as well as based on the SNPs within the top 20% with regard to their p-value. Since roughly 2100 genes are covered by the IBC chip, a p-value $< 2.4 \times 10^{-5}$ is considered significant.

2.5. Replication

Given that contemporary genetics consortia and results indicate that very large replication samples are needed to successfully replicate SNPs, and that false positives are an ongoing issue in studies such as these, we attempted to replicate only the two SNPs that met experiment-wide significance for ABI in European Americans in an additional 13,524 individuals of European ancestry from six population-based studies (Copenhagen city heart study, $n = 5182$; genetic study of aspirin responsiveness (GeneSTAR), $n = 618$; KORA F3 (independent of KORA F3 participants in the discovery sample), $n = 1440$; KORA F4, $n = 411$; national health and nutrition examination survey (NHANES), $n = 2358$, and prevention of renal and vascular end-stage disease (PREVEND), $n = 3515$) and 1916 individuals of European ancestry from clinically based samples (cardiovascular disease in intermittent claudication (CAVASIC), $n = 434$; genetic determinants of peripheral arterial disease (GenePAD), $n = 811$; and Linz peripheral arterial disease

Table 2A
Candidate gene SNP associations for ankle-brachial index: discovery meta-analysis, $p < 10^{-4}$. European and European Americans, $n = 21{,}547$.

SNP	Chr	Physical position	Nearest gene	Feature	Major/minor allele	MAF[a]	Beta	95% CI	p Value	p_{het}[b]
rs2171209	6	159,103,550	*SYTL3*	Intron	C/T	0.22	−0.007	−0.010, −0.004	6.02×10^{-7}	0.55
rs290481	10	114,913,815	*TCF7L2*	Intron	C/T	0.17	−0.008	−0.011, −0.005	7.01×10^{-7}	0.08
rs11061318	12	130,135,058	*GPR133*	Missense Ser → Leu	C/T	0.03	−0.016	−0.023, −0.009	4.46×10^{-6}	0.91

[a] MAF = minor allele frequency.
[b] p-for-heterogeneity by Cochran's Q.

(LIPAD), $n = 671$). GeneSTAR provided in silico genotyping (genotyped participants with the same candidate gene chip used in the discovery cohorts) while in the remaining studies genotyped the 2 SNPs de novo using Taqman or Sequenom genotyping platforms. Description of the replication studies, ABI protocol and calculation, and participant characteristics are provided in Supplementary Methods, Supplementary Method Table 1 and Supplementary Results Table 1.

2.6. Power and sample size for discovery and replication stages

Participants of European ancestry: With a minor allele frequency (MAF) of 0.10, additive SNP modeling, and experiment-wide significance level of 2×10^{-6}, for each additional copy of the risk allele, we have 80% power to detect a beta coefficient for ABI of 0.0092 and an OR for PAD of 1.45 in the discovery stage ($n = 21{,}547$). For European ancestry individuals in the replication stage for ABI ($n = 15{,}440$) using the more stringent significance level of 2×10^{-6} we have 80% power to detect a beta coefficient of 0.0106; using a less stringent level (i.e. $\alpha = 0.025$ based on carrying forward 2 SNPs for replication), we have 80% power to detect a beta coefficient of 0.0058. In the discovery plus replication stage for ABI ($n = 36, 987$), we have 80% power to detect a beta coefficient of 0.0065 using 2×10^{-6} as the type 1 error rate.

Participants of African-American ancestry: With a minor allele frequency (MAF) of 0.10, additive SNP modeling, and experiment-wide significance level of 2×10^{-6}, for each additional copy of the risk allele, we have 80% power to detect a beta coefficient for ABI of 0.0155 and an OR for PAD of 1.67 in the sample of 7267 individuals in the discovery stage. No replication sample was available for this ethnicity.

3. Results

3.1. European ancestry studies: meta-analysis of ABI and PAD

In European ancestry discovery samples, two SNPs were significantly associated with ABI (Table 2A, Figs. 1 and 2): each additional copy of the minor allele of rs2171209 in *SYTL3* was associated with a 0.007 lower ABI (95% CI −0.010, −0.004, $p = 6.02 \times 10^{-7}$, p for heterogeneity = 0.55) and each additional copy of the minor allele of rs290481 in *TCF7L2* was associated with a 0.008 lower ABI (95% CI −0.011, −0.005, $p = 7.01 \times 10^{-7}$, p for heterogeneity = 0.08). Rs290481 is located in intron 14 within the 3′ region of the *TCF7L2* gene on chromosome 10 is distinct from a cluster of SNPs in the 5′ region of *TCF7L2* (represented by rs7903146), previously reported to be associated with type 2 diabetes in genome-wide association studies ($r^2 = 0.001$ between rs290481 and rs7903146) [21]. Among those of European ancestry, rs2171209 was not significantly associated with the categorical PAD diagnosis made by the ABI threshold of 0.90 – each additional copy of the minor allele was associated with just a 1.09-fold greater odds of PAD (95% CI: 0.97, 1.22, $p = 0.14$); however, rs290481 was associated with PAD (OR = 1.20, 95% CI: 1.06, 1.35, $p = 0.004$), although this

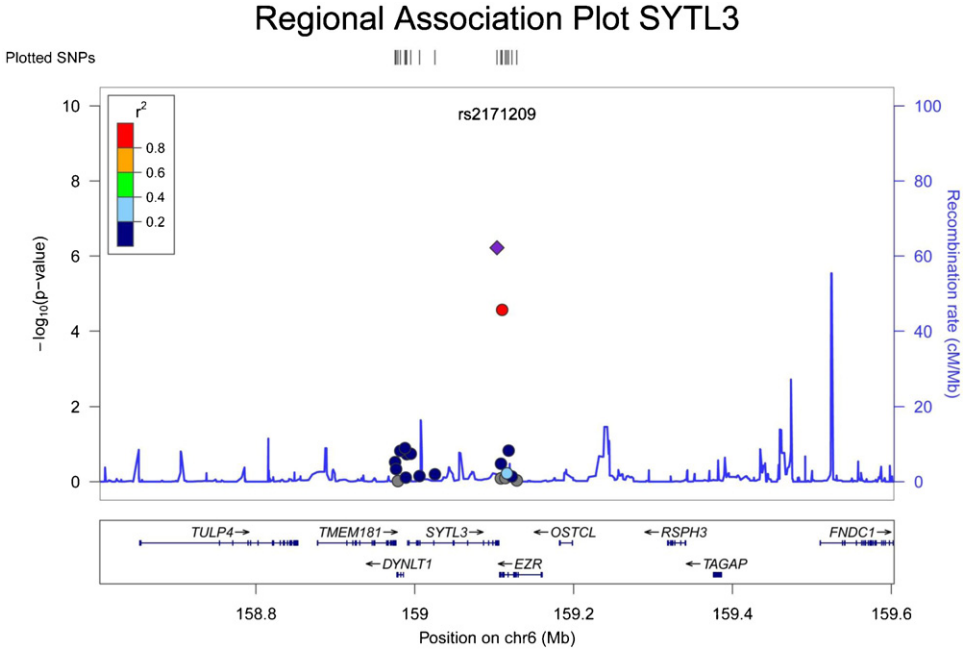


Fig. 1. This plot shows the p-values for rs2171209, as well as for SNPs in the region of rs2171209, with ABI in a meta-analysis of the discovery studies. The x-axis shows chromosomal location in Mb (chromosome 10), as well as genes residing in this region. The y-axis on the left displays the −log 10(p-value) for each SNP, and the y-axis on the right shows the recombination rate in this region. The top SNP, rs2171209 is represented as a purple diamond, while supporting SNPs and other SNPs in the area are color-coded by linkage disequilibrium with rs2171209 (see r^2 linkage disequilibrium legend on the plot). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

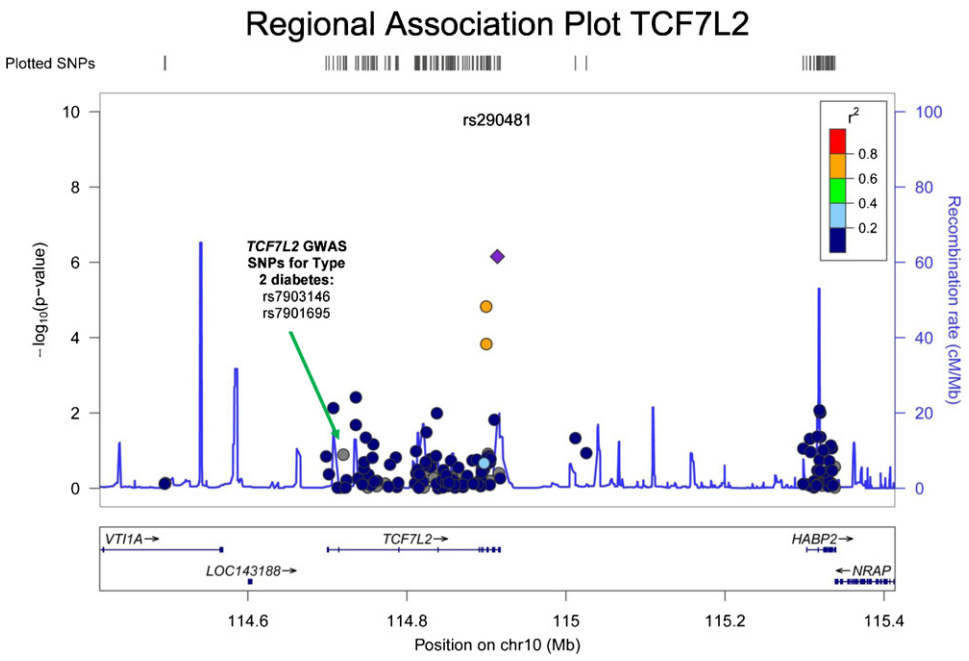


Fig. 2. This plot shows the *p*-values for rs290481, as well as for SNPs in the region of rs290481, with ABI in a meta-analysis of the discovery studies. The *x*-axis shows chromosomal location in Mb (chromosome 10), as well as genes residing in this region. The *y*-axis on the left displays the −log 10(*p*-value) for each SNP, and the *y*-axis on the right shows the recombination rate in this region. The top SNP, rs290481 is represented as a purple diamond, while supporting SNPs and other SNPs in the area are color-coded by linkage disequilibrium with rs290481 (see r^2 linkage disequilibrium legend on the plot). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

association did not meet the experiment-wide significance level. There was no association between rs290481 and ABI ($\beta = -0.001$ 95% CI: −0.006, 0.005, $p = 0.80$) or rs2171209 and ABI ($\beta = -0.002$ 95% CI −0.009, 0.004 $p = 0.46$) detected in African Americans. One additional SNP was nominally associated with ABI in individuals of European ancestry at $p < 10^{-4}$ (Table 2A). The most significant SNP associations for ABI in models minimally adjusted for age, sex, and study site only remained similar to the fully adjusted models. When we removed all participants with type 2 diabetes from the analysis of the CARe studies (ARIC, CFS, CHS, FHS and MESA), the beta coefficients were essentially unchanged from those in Table 2A. Similarly, when we removed participants less than 60 years of age from the analysis in the CARe studies, beta coefficients were essentially unchanged from those in Table 2A.

Because rs290481 is located in a gene previously strongly associated with type 2 diabetes, we conducted a test for interaction of this genotype with type 2 diabetes for ABI within the CARe discovery cohorts (ARIC, CFS, CHS, FHS, MESA). The test for interaction was performed on an additive scale using linear regression within each CARe cohort, and then combining the results in the diabetes and no diabetes strata using fixed effect inverse-variance weighting meta-analysis. The magnitude of the effect with ABI was greater in diabetics (*p* for interaction 0.04, Supplementary Results Table 2, with each copy of the minor allele in participants with diabetes conferring a lower level of ABI ($n = 1896$, $\beta = -0.03$, $p < 0.001$) compared to participants without diabetes ($n = 16{,}685$, $\beta = -0.007$, $p = 0.007$).

The gene-based analysis using VEGAS at the discovery stage did not reveal any different significantly associated genes with continuous ABI from our individual SNP analysis (Supplementary Results Table 3). *SYTL3* and *TCF7L2* were among the most significant genes for both analyses using all SNPs and the top 20% of SNPs (*SYTL3*: *p*(all SNPs) = 0.00216, *p*(top20% of SNPs) = 0.00007; *TCF7L2*: *p*(all SNPs) = 0.03589 *p*(top20% of SNPs) = 0.00161).

In the replication stage, the association between ABI and rs2171209 in *SYTL3* was not significant in the population-based ($\beta = -0.0004$, $n = 13{,}510$, $p = 0.73$) and clinically based replication samples ($\beta = 0.001$, $n = 1890$, $p = 0.82$) (Fig. 3). Consequently, in the combined discovery plus replication meta-analysis the association was no longer significant ($n = 36{,}947$, β −0.003, $p = 1.14 \times 10^{-3}$) (Fig. 3). rs290481 in *TCF7L2* also failed to replicate in the population-based replication studies ($\beta = -0.001$, $n = 13{,}505$, $p = 0.38$) and in the clinically based replication studies ($\beta = -0.008$, $n = 1896$, $p = 0.20$) (Fig. 4). In the combined discovery and replication meta-analysis the association between rs290481 and ABI no longer met experiment-wide significance ($n = 36{,}855$, β −0.004, $p = 8.88 \times 10^{-5}$) (Fig. 4).

None of the SNP associations in individuals of European ancestry achieved experiment-wide significance for PAD (Table 2B). One of the most significant associations for PAD was in a coding, non-synonymous SNP rs3745274 on chromosome 19 in *CYP2B6* (OR 1.24, $p = 4.99 \times 10^{-5}$).

3.2. African-Americans: meta-analysis of ABI and PAD

In African Americans none of the SNP associations with ABI and PAD were statistically significant (Tables 3A and 3B). The strongest association for ABI was rs2243100 on chromosome 17 in *SLC25A11* (β 0.011 95% CI 0.006, 0.017, $p = 5 \times 10^{-5}$) and for PAD was rs4987756 on chromosome 18 in *BCL2* (OR 2.99, 95% CI 1.88, 4.76, $p = 3.78 \times 10^{-6}$).

4. Discussion

We conducted a large candidate gene association study of ~2100 cardiovascular candidate genes for ABI and PAD in over 21,000 individuals of European ancestry and over 7000 African Americans. In individuals of European ancestry, a SNP in the *TCF7L2* gene (rs290481) and a SNP in the *SYTL3* gene (rs2171209) were significantly associated in the discovery stage with variation in ABI measurements and a suggestive association was identified in a SNP in *CYP2B6* for PAD. These findings are intriguing as the genes are linked to key PAD risk factors. *TCF7L2* is the strongest

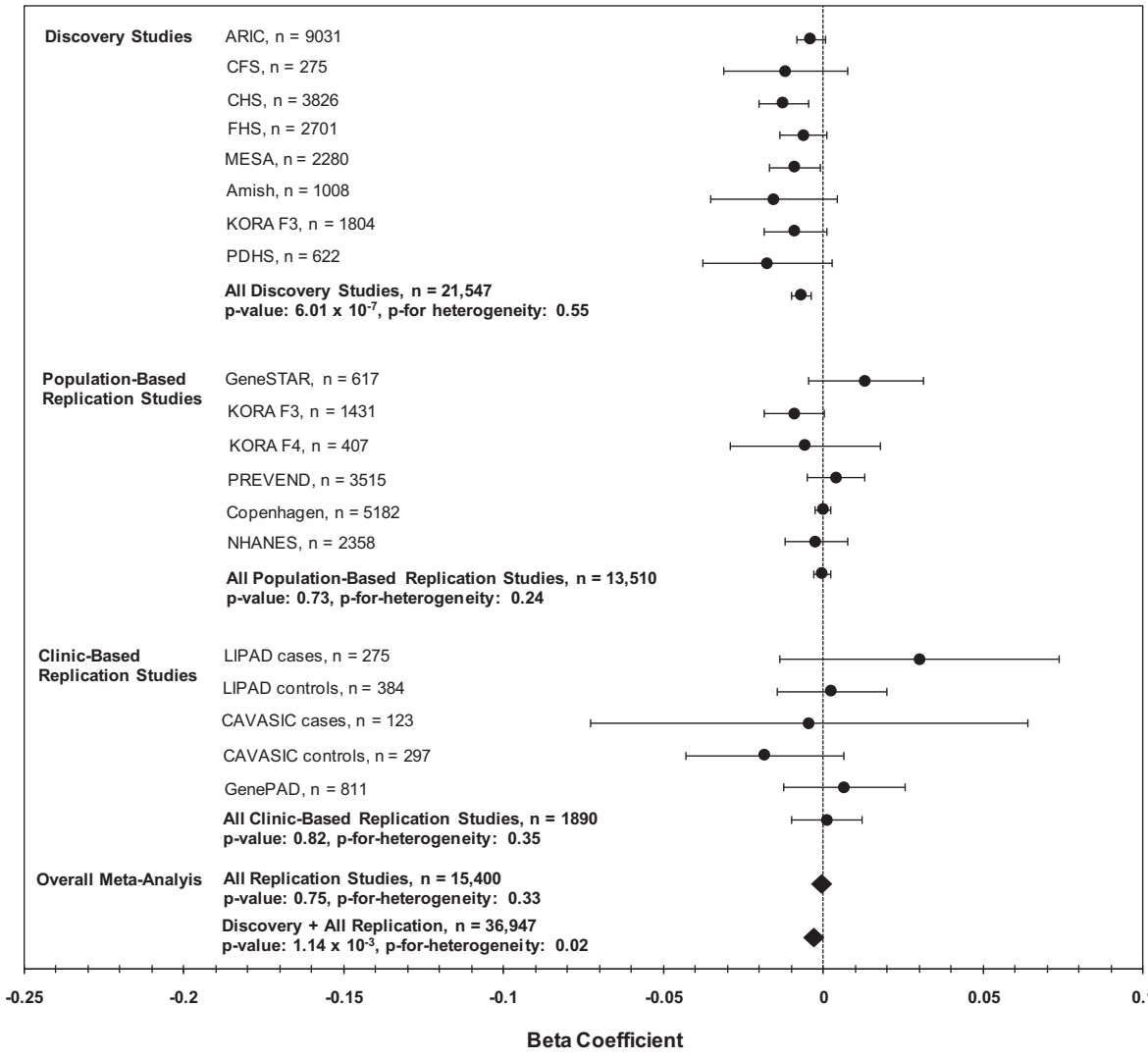


Fig. 3. This plot shows the association of rs2171209 with ABI for each of the discovery and replication studies, and then results with these studies combined by meta-analysis. The *x*-axis contains beta coefficients for the association of rs2171209 with ABI in the discovery and replication studies, and the *y*-axis shows studies or groups of studies. Circles represent the beta coefficient (except for overall replication and discovery + replication meta-analysis results, where beta coefficients are designated by a diamond), and error bars are 95% confidence intervals. *p*-Values for heterogeneity are by Cochran's Q. Study abbreviations are as follows: ARIC = atherosclerosis risk in communities, CFS = Cleveland family study, CHS = cardiovascular health study, FHS = Framingham heart study, MESA = multi-ethnic study of atherosclerosis, KORA F3 and KORA F4 = cooperative research in the region of Augsburg, PDHS = Penn diabetes heart study, GeneSTAR = genetic study of aspirin responsiveness, PREVEND = prevention of renal and vascular end-stage disease, Copenhagen = Copenhagen city heart study, NHANES = national health and nutrition examination survey, LIPAD = Linz peripheral arterial disease, CAVASIC = cardiovascular disease in intermittent claudication, and GenePAD = genetic determinants of peripheral arterial disease.

Table 2B
Candidate gene SNP associations for PAD (ABI < 0.9): discovery meta-analysis, $p < 10^{-4}$. European and European Americans, $n = 20{,}539$.

SNP	Chr	Physical position	Nearest gene	Feature	Major/minor allele	MAF[a]	Odds ratio	95% CI	*p* Value	p_{het}[b]
rs11088283	21	34,745,649	*KCNE1*	Intron	A/G	0.47	0.85	(0.78, 0.93)	4.88×10^{-5}	0.09
rs3745274	19	46,204,681	*CYP2B6*	Coding, non-synonymous	G/T	0.24	1.24	(1.12, 1.38)	4.99×10^{-5}	0.01
rs12428227	13	109,700,293	*COL4A1*	Intron	A/G	0.17	1.23	(1.10, 1.38)	5.20×10^{-5}	0.14
rs17151901	8	10,290,865	*MSRA*	Intron	C/T	0.02	1.36	(1.09, 1.69)	6.53×10^{-5}	0.01

[a] MAF = minor allele frequency.
[b] *p*-for-heterogeneity by Cochran's Q.

Table 3A
Candidate gene SNP associations for ankle-brachial index: discovery meta-analysis, $p < 10^{-4}$. African Americans, $n = 7267$.

SNP	Chr	Physical position	Nearest gene	Feature	Major/minor allele	MAF[a]	Beta	95% CI	*p* Value	p_{het}[b]
rs2243100	17	4,779,777	*SCL25A11*	Locus-region	C/T	0.20	0.011	0.006, 0.017	5.00×10^{-5}	0.32
rs2243093	17	4,776,675	*GP1BA*	5′ UTR	T/C	0.22	0.012	0.007, 0.017	5.94×10^{-5}	0.42
rs2660896	12	94,947,913	*LTA4H*	Intron	C/T	0.25	−0.010	−0.015, −0.005	7.94×10^{-5}	0.96
rs2242406	16	74,131,531	*CHST5*	Intron	C/T	0.01	−0.034	−0.051, −0.017	9.46×10^{-5}	0.18

[a] MAF = minor allele frequency.
[b] *p*-for-heterogeneity by Cochran's Q.

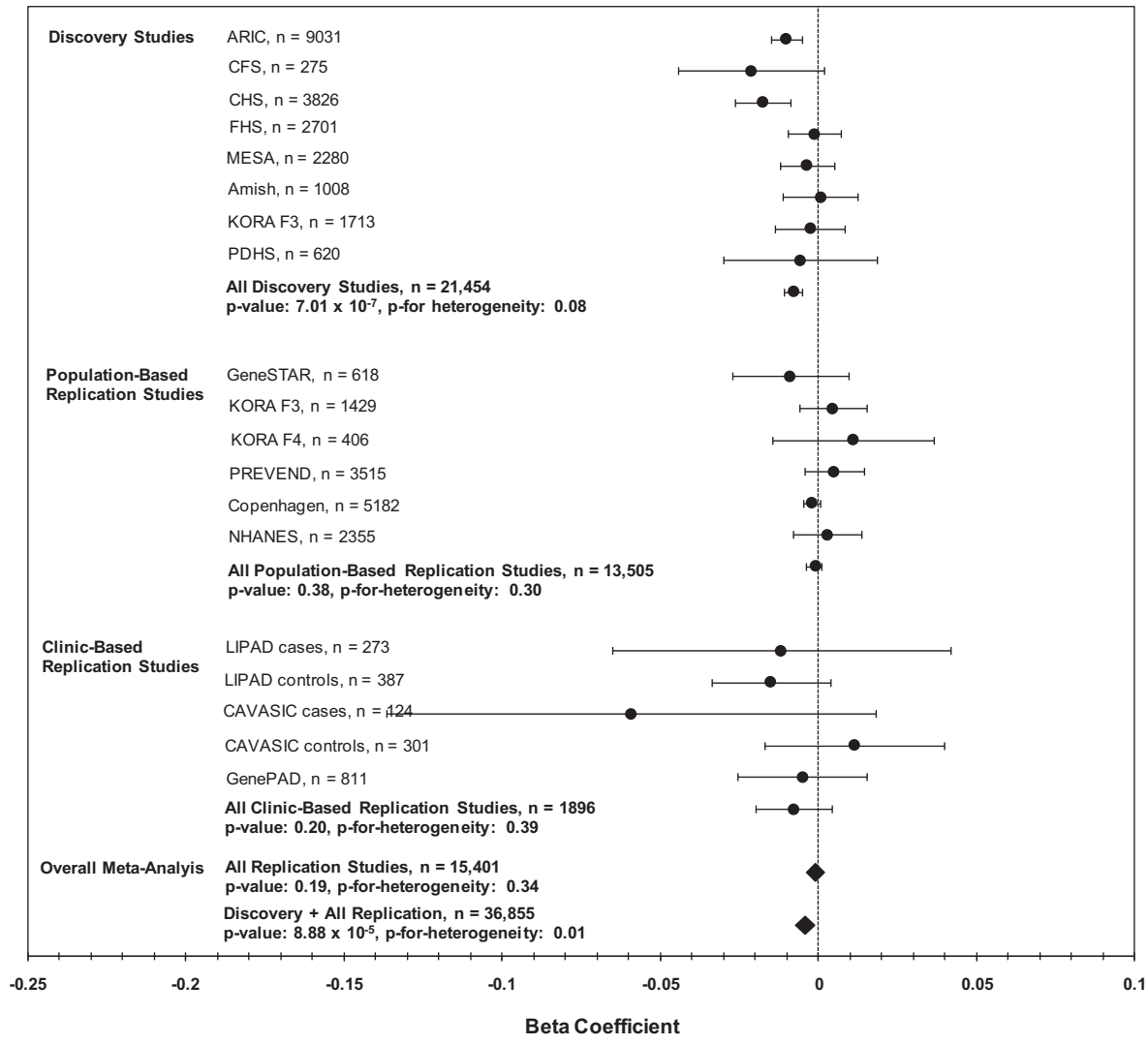


Fig. 4. This plot shows the association of rs290481 with ABI for each of the discovery and replication studies, and then results with these studies combined by meta-analysis. The *x*-axis contains beta coefficients for the association of rs290481 with ABI in the discovery and replication studies, and the *y*-axis shows studies or groups of studies. Circles represent the beta coefficient (except for overall replication and discovery + replication meta-analysis results, where beta coefficients are designated by a diamond), and error bars are 95% confidence intervals. *p*-Values for heterogeneity are by Cochran's Q. Study abbreviations are as follows: ARIC = atherosclerosis risk in communities, CFS = Cleveland family study, CHS = cardiovascular health study, FHS = Framingham heart study, MESA = multi-ethnic study of atherosclerosis, KORA F3 and KORA F4 = cooperative research in the region of Augsburg, PDHS = Penn diabetes heart study, GeneSTAR = genetic study of aspirin responsiveness, PREVEND = prevention of renal and vascular end-stage disease, Copenhagen = Copenhagen city heart study, NHANES = national health and nutrition examination survey, LIPAD = Linz peripheral arterial disease, CAVASIC = cardiovascular disease in intermittent claudication, and GenePAD = genetic determinants of peripheral arterial disease.

Table 3B
Candidate gene SNP associations for PAD (ABI < 0.9): discovery meta-analysis, $p < 10^{-4}$. African Americans, $n = 7267$.

SNP	Chr	Physical position	Nearest gene	Feature	Major/minor allele	MAF[a]	Odds ratio	95% CI	*p* Value	p_{het}[b]
rs4987756	18	59,060,091	*BCL2*	Intron	A/G	0.01	2.99	1.88, 4.76	3.78×10^{-6}	0.63
rs1256143	14	63,981,380	*MTHFD1*	Intron	C/T	0.19	1.72	1.21, 1.66	1.43×10^{-5}	0.05
rs13004470	2	242,159,756	*BOK*	Intron	C/T	0.18	1.39	1.19, 1.62	4.69×10^{-5}	0.81
rs9830448	3	154,349,978	*RAPB2*	Locus region	C/A	0.07	1.58	1.28, 2.02	4.79×10^{-5}	0.73

[a] MAF = minor allele frequency.
[b] *p*-for-heterogeneity by Cochran's Q.

genetic risk factor for susceptibility to type 2 diabetes [22–25] and *CYP2B6* affects smoking behavior [26] and thus may be important in tobacco-related diseases such as PAD. However, we were unable to replicate the SNP–ABI associations in additional samples from population-based studies or clinically based samples. Furthermore, the associations were not detected in African Americans. We did not observe any significant associations for ABI or PAD in African Americans, possibly due to the relatively small sample size limiting our power to detect associations.

4.1. In the context of the current literature

Genetic factors leading to susceptibility to PAD remain largely unknown but are likely to be attributed to variants in many genes, each with small effects [13] or possibly from rare variants (minor allele frequency < 1%) with larger effects. While many of these variants may lead to risk for PAD through effects on established risk factors or shared effects with CAD and other atherosclerotic diseases [27], other variants may uniquely influence development

of arterial disease in the lower extremities. Although our findings after the discovery stage did not bear out in the replication samples, the two genes are interesting candidates for ABI in light of the current literature and deserve some discussion. Genome-wide association studies of individuals of European ancestry have consistently reported an association between genetic variants of *TCF7L2* and type 2 diabetes that has been confirmed in Japanese and African American samples [19,22–24,28–30]. However, our SNP in *TCF7L2*, rs290481 was not significantly associated with type 2 diabetes in large scale association analysis [21]. The replicated index SNP in *TCF7L2* associated with type 2 diabetes is rs7902146. Further, rs290481 is not in linkage disequilibrium with rs7902146 ($r^2 = 0.001$) [21]. *TCF7L2* encodes a high mobility group (HMG) box-containing transcription factor that is involved in the Wnt signaling pathway [31] and is associated with impaired beta cell function, impaired insulin secretion and increased hepatic glucose production. Therefore, the *TCF7L2* association in our discovery cohorts that presented even after adjusting for type 2 diabetes might deserve further attention in functional studies to elucidate its role in atherosclerosis.

The association between *SYTL3* and ABI may be mediated by lipoprotein(a) (Lp(a)). A genome-wide association study in a small founder population of 386 Hutterites identified an association between the extended *LPA* gene region on chromosome 6q26–q27 including *SYTL3* SNPs with Lp(a) levels[32]. Genetic variation within the *LPA* gene region including a very common copy number variation and other genetic variants explain up to 90% of Lp(a) concentrations[33]. Since Lp(a) concentrations and genetic variants within the *LPA* region are a strong risk factor for cardiovascular disease[34], it might well be that a SNP in the *SYTL3* gene reflects a signal from *LPA*. Polymorphisms within the *LPA* gene region were associated with PAD in a past study [35]. Lp(a) may be an independent risk factor for PAD [36] but results are conflicting [37,38].

4.2. Strengths and limitations

To our knowledge this study is the largest candidate gene association study concerning ABI conducted in both individuals of European ancestry and African Americans and includes the most extensive number of candidate genes investigated. In the CARe consortium, imputed GWAS data is available on the African-American participants; however, given our relatively modest African-American sample size, we chose to perform analysis of the IBC chip first. We have also chosen the IBC chip because it was specifically designed as a large scale cardiovascular-centric candidate gene array, and the genetic variants on the chip were informed by GWAS for vascular and inflammatory diseases as well as expression QTLs for atherosclerosis.

Several limitations of our candidate gene meta-analysis merit comment: (i) The ankle-brachial blood pressure measurement protocols used in the studies were heterogeneous. Hence, phenotype heterogeneity may have impacted our ability to detect associations. (ii) The ARIC study contributed over 40% of the European ancestry sample and measured ABI in only one leg which may have led to phenotype misclassification most problematic for the PAD phenotype. The mean ABI did differ significantly between European Americans in ARIC and European Americans from the other CARe cohorts (all $p < 0.05$) with mean ± SD of the ARIC ABI 1.12 ± 0.13, CHS ABI 1.06 ± 0.15, CFS ABI 1.08 ± 0.10, FHS ABI 1.13 ± 0.12, and MESA ABI 1.11 ± 0.12. However, a sensitivity analysis excluding the ARIC samples showed parameter estimates of similar size. (iii) Not all studies had information on lower extremity revascularization, which may also have contributed to PAD misclassification. In general, these misclassifications should cause bias toward the null. (iv) Control selection bias could have affected our PAD results in some way, although given that all of our studies except one (the PDHS) contributing to the PAD analysis were prospective cohort studies where knowledge of PAD would not affect exposure (i.e. genotype status) and the genotype precedes prevalent PAD, this is of lesser concern. For ABI analyses, we also analyzed our clinical replication samples separately by case-control status to avoid additional bias or heterogeneity. (v) Although we adjusted for population stratification using principal components in the European ancestry analysis and global ancestry in the African-American analysis, residual confounding could still be present. (vi) Our sample of African Americans was modest in size and likely limited our power to detect associations. For example, given the sample size of African-Americans we included and a risk allele frequency of 0.10, we only had 80% power to detect an increment in ABI of approximately 0.02 or greater per each copy of the risk allele. The observed effect size of the experiment-wide significant SNPs in European ancestry participants was much smaller than this value. According to our calculations in the methods section, for the European ancestry analyses, we can detect modest differences in ABI (similar to the ones we observed in this study), but are likely underpowered for PAD.

Some of the mentioned limitations might have contributed to the observation that the most important findings from the discovery phase could not be confirmed in the replication phase. However, it is unlikely that this fully explains the differences between the two study stages which necessitate additional large study samples.

5. Conclusions

The search for genes influencing ABI and PAD remains challenging. Although we cannot claim new findings in our study, two associations at the discovery stage for ABI (*SYTL3*, *TCF7L2*) may deserve further attention in other populations and functional studies. Further study of the genes identified in this study for ABI (*SYTL3*, *TCF7L2*) and PAD (*CYP2B6*) is warranted in other populations as further investigation of the function of these loci may uncover important biological insights into the pathogenesis of PAD. Identification of main effects may have been difficult in our study due to the presence of interactions and heterogeneity across participating studies. New and more powerful approaches to PAD gene discovery are sorely needed.

Funding

The Candidate Gene Association Resource (CARe) is supported by contract number HHSN268200625226C from the National Institutes of Health (NIH)/National Heart Lung and Blood Institute (NHLBI), and subcontract number 5215810-55000000041 to C.L.W. A full listing of the grants and contracts that have supported CARe is provided at http://public.nhlbi.nih.gov/GeneticsGenomics/home/care.aspx. Please see information in supplementary information file for a complete list of funding information for each study participating in this manuscript.

Acknowledgements

The authors would like to acknowledge the participants of all studies in the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.atherosclerosis.2012.01.039.

References

[1] Ankle Brachial Index Collaboration FG, Fowkes GD, Murray I, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. JAMA 2008;300:197–208.

[2] Steg PG, Bhatt DL, Wilson PW, et al. One-year cardiovascular event rates in outpatients with atherothrombosis. JAMA 2007;297:1197–206.

[3] Bhatt DL, Steg PG, Ohman EM, et al. International prevalence, recognition, and treatment of cardiovascular risk factors in outpatients with atherothrombosis. JAMA 2006;295:180–9.

[4] Hirsch AT, Haskal ZJ, Hertzer NR, et al. ACC/AHA 2005 Practice Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease): endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation. Circulation 2006;113:e463–654.

[5] Valentine RJ, Guerra R, Stephan P, et al. Family history is a major determinant of subclinical peripheral arterial disease in young adults. J Vasc Surg 2004;39:351–6.

[6] Valentine RJ, Verstraete R, Clagett GP, Cohen JC. Premature cardiovascular disease is common in relatives of patients with premature peripheral atherosclerosis. Arch Intern Med 2000;160:1343–8.

[7] Carmelli D, Fabsitz RR, Swan GE, et al. Contribution of genetic and environmental influences to ankle-brachial blood pressure index in the NHLBI Twin Study. National Heart, Lung, and Blood Institute. Am J Epidemiol 2000;151: 452–8.

[8] Kullo IJ, Turner ST, Kardia SL, et al. A genome-wide linkage scan for ankle-brachial index in African American and non-Hispanic white subjects participating in the GENOA study. Atherosclerosis 2006;187:433–8.

[9] Murabito JM, Guo CY, Fox CS, D'Agostino RB. Heritability of the ankle-brachial index: the Framingham Offspring study. Am J Epidemiol 2006;164:963–8.

[10] Wassel CL, Loomba R, Ix JH, et al. Family history of peripheral artery disease is associated with prevalence and severity of peripheral artery disease: the San Diego population study. J Am Coll Cardiol 2011;58:1386–92.

[11] Newman AB, Shemanski L, Manolio TA, et al. Ankle-arm index as a predictor of cardiovascular disease and mortality in the Cardiovascular Health Study. The Cardiovascular Health Study Group. Arterioscler Thromb Vasc Biol 1999;19:538–45.

[12] Murabito JM, White CC, Kavousi M, et al. Association between chromosome 9p21 variants and the ankle-brachial index identified by a meta-analysis of 21 genome-wide association studies. Circ Cardiovasc Genet 2011. Epub ahead of print.

[13] Knowles JW, Assimes TL, Li J, Quertermous T, Cooke JP. Genetic susceptibility to peripheral arterial disease: a dark corner in vascular biology. Arterioscler Thromb Vasc Biol 2007;27:2068–78.

[14] Keating BJ, Tischfield S, Murray SS, et al. Concept, design and implementation of a cardiovascular gene-centric 50K SNP array for large-scale genomic association studies. PLoS ONE 2008;3:e3583.

[15] Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med 2008;358:1240–9.

[16] De Caterina R, Talmud PJ, Merlini PA, et al. Strong association of the APOA5-1131T>C gene variant and early-onset acute myocardial infarction. Atherosclerosis 2011;214:397–403.

[17] Musunuru K, Lettre G, Young T, et al. Candidate gene association resource (CARe): design, methods, and proof of concept. Circ Cardiovasc Genet 2010;3:267–75.

[18] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.

[19] Lettre G, Palmer CD, Young T, et al. Genome-wide association study of coronary heart disease and its risk factors in 8090 African Americans: the NHLBI CARe Project. PLoS Genet 2011;7:e1001300.

[20] Liu JZ, McRae AF, Nyholt DR, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet 2010;87:139–45.

[21] Voight BF, Scott LJ, Steinthorsdottir V, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–89.

[22] Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336–41.

[23] Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–5.

[24] Saxena R, Gianniny L, Burtt NP, et al. Common single nucleotide polymorphisms in TCF7L2 are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals. Diabetes 2006;55:2890–5.

[25] Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, Novartis Institutes of BioMedical Research, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331–6.

[26] Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452:638–42.

[27] Cluett C, McDermott MM, Guralnik J, et al. The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people. Circ Cardiovasc Genet 2009;2:347–53.

[28] Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–5.

[29] Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007;39:770–5.

[30] Takeuchi F, Serizawa M, Yamamoto K, et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. Diabetes 2009;58:1690–9.

[31] Jin T, Liu L. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. Mol Endocrinol 2008;22:2383–92.

[32] Ober C, Nord AS, Thompson EE, et al. Genome-wide association study of plasma lipoprotein(a) levels identifies multiple genes on chromosome 6q. J Lipid Res 2009;50:798–806.

[33] Utermann G. Genetic architecture and evolution of the lipoprotein(a) trait. Curr Opin Lipidol 1999;10:133–41.

[34] Erqou S, Thompson A, Di Angelantonio E, et al. Apolipoprotein(a) isoforms and the risk of vascular disease: systematic review of 40 studies involving 58 000 participants. J Am Coll Cardiol 2010;55:2160–7.

[35] Dieplinger B, Lingenhel A, Baumgartner N, et al. Increased serum lipoprotein(a) concentrations and low molecular weight phenotypes of apolipoprotein(a) are associated with symptomatic peripheral arterial disease. Clin Chem 2007;53:1298–305.

[36] Volpato S, Vigna GB, McDermott MM, et al. Lipoprotein(a), inflammation, and peripheral arterial disease in a community-based sample of older men and women (the InCHIANTI study). Am J Cardiol 2010;105:1825–30.

[37] Pradhan AD, Shrivastava S, Cook NR, et al. Symptomatic peripheral arterial disease in women: nontraditional biomarkers of elevated risk. Circulation 2008;117:823–31.

[38] Molgaard J, Klausen IC, Lassvik C, et al. Significant association between low-molecular-weight apolipoprotein(a) isoforms and intermittent claudication. Arterioscler Thromb 1992;12:895–901.

BMC Public Health

RESEARCH ARTICLE Open Access

Randomised controlled trial of an automated, interactive telephone intervention (TLC Diabetes) to improve type 2 diabetes management: baseline findings and six-month outcomes

Emily D Williams[1*], Dominique Bird[1], Andrew W Forbes[1], Anthony Russell[2,3], Susan Ash[4], Robert Friedman[5], Paul A Scuffham[6] and Brian Oldenburg[1]

Abstract

Background: Effective self-management of diabetes is essential for the reduction of diabetes-related complications, as global rates of diabetes escalate.

Methods: Randomised controlled trial. Adults with type 2 diabetes (n = 120), with HbA_{1c} greater than or equal to 7.5 %, were randomly allocated (4 × 4 block randomised block design) to receive an automated, interactive telephone-delivered management intervention or usual routine care. Baseline sociodemographic, behavioural and medical history data were collected by self-administered questionnaires and biological data were obtained during hospital appointments. Health-related quality of life (HRQL) was measured using the SF-36.

Results: The mean age of participants was 57.4 (SD 8.3), 63% of whom were male. There were no differences in demographic, socioeconomic and behavioural variables between the study arms at baseline. Over the six-month period from baseline, participants receiving the Australian TLC (Telephone-Linked Care) Diabetes program showed a 0.8% decrease in geometric mean HbA_{1c} from 8.7% to 7.9%, compared with a 0.2% HbA_{1c} reduction (8.9% to 8.7%) in the usual care arm (p = 0.002). There was also a significant improvement in mental HRQL, with a mean increase of 1.9 in the intervention arm, while the usual care arm decreased by 0.8 (p = 0.007). No significant improvements in physical HRQL were observed.

Conclusions: These analyses indicate the efficacy of the Australian TLC Diabetes program with clinically significant post-intervention improvements in both glycaemic control and mental HRQL. These observed improvements, if supported and maintained by an ongoing program such as this, could significantly reduce diabetes-related complications in the longer term. Given the accessibility and feasibility of this kind of program, it has strong potential for providing effective, ongoing support to many individuals with diabetes in the future.

Background

The rapid increase in rates of diabetes poses a significant public health problem globally. Diabetes is currently estimated to affect 285 million adults worldwide, with the prevalence predicted to rise to 438 million by the year 2030 [1]. Its complications contribute significantly to ill health, disability, poor quality of life and premature death. The associated global economic burden is projected to reach at least US$376 billion in 2030 [2]. Although guidelines and targets for optimal diabetes management are well documented [3], it is estimated that 40% of individuals with diabetes have sub-optimal glycaemic control [4,5], significantly increasing their risk of costly and debilitating diabetes-related complications [6,7].

Diabetes self-management education facilitates the acquisition of knowledge and skills to improve disease management and has been found to improve glycaemic control [8], with program duration being a critical

* Correspondence: emily.d.williams@monash.edu
[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
Full list of author information is available at the end of the article

predictor of this success [9]. Providing ongoing and long-term diabetes management support, particularly to those people living in rural and remote areas, is a major challenge for all health systems around the world. This highlights the need to develop and evaluate more feasible, accessible ways of providing such support for large numbers of people with diabetes than is traditionally offered. Using information and communication technology (ICT) to provide diabetes management education and support directly to patients offers such potential, by overcoming many of the barriers associated with more traditional modes of program delivery. Use of ICT has been shown to yield improvements in self-care knowledge and behaviour of patients and clinical outcomes associated with the prevention and control of chronic health conditions, including diabetes [10-12]. Some studies have evaluated the role of automated or semi-automated telephone-delivered diabetes management interventions on glycaemic control, however, the results have been inconsistent with varying levels of reliance upon health professionals [13-15].

The Telephone-Linked Care (TLC) program is an automated and interactive telephone system designed to emulate telephone encounters between patients and health professionals [16] and to complement standard medical care. TLC systems have been previously used to effectively screen people with specific health conditions [17,18], promote self-care behaviours [19-22] and provide monitoring of and feedback to patients with a range of chronic diseases [23-26].

A randomised controlled trial was conducted to evaluate a TLC program - the Australian TLC Diabetes program - designed to improve type 2 diabetes management. This paper presents the six-month results for the study's primary outcomes, glycosylated haemoglobin and health-related quality of life (HRQL), and it also describes the sample baseline characteristics, compared with a large Australian population study.

Methods

Study design

The study methodology has been detailed elsewhere [27]. In brief, the study was a two-arm prospective randomised controlled trial, with adults with type 2 diabetes randomised to either the intervention (Australian TLC Diabetes program) arm or 'usual care' control arm. Data were collected between July 2008 and December 2010. Ethics approval was received from the Human Research Ethics Committees for all collaborating hospitals and Monash University.

Participant recruitment and randomisation

Participants were recruited through advertisements in newspapers, flyers distributed to health professionals and to members of Diabetes Australia – Queensland, community newsletters and through diabetes clinics at three major hospitals in Brisbane (Princess Alexandra Hospital, Royal Brisbane and Women's Hospital, and Prince Charles Hospital).

There were two steps to the eligibility screening (Table 1). In the first step, which took place during the initial contact via telephone or in person, research staff excluded individuals who did not meet all of the Step 1 eligibility criteria or who met any of the Step 1 exclusion criteria. If potentially eligible, participants attended a baseline appointment at either Princess Alexandra or Royal Brisbane and Women's Hospital, where full information was provided, informed consent was obtained and baseline data collected. At that appointment, baseline questionnaires were completed and fasting blood specimens were taken, along with other clinical data (blood pressure, weight, height and waist circumference). Blood tests were conducted by Queensland Pathology using standardised assays. The second screening step verified the glycosylated haemoglobin (HbA_{1c}) inclusion criterion (≥ 7.5%). The final sample included 120 adults; n = 60 in each of the study arms. The allocation ratio was 1:1 and the allocation sequence was computer-generated. The arm allocation was conducted using a 4x4 block randomised block design with the participant as the unit of randomisation.

Study arms

All participants received a quarterly newsletter containing general health information; this aimed to maintain participation in both arms. Participants in both arms were advised to continue with their usual medical care. The usual care arm received no further intervention. The treating physicians were not blinded to the allocation.

Intervention arm

Australian TLC Diabetes program

The intervention took place over six months during which they received the Australian TLC Diabetes program. Its main component is the Telephone-Linked Care (TLC) Diabetes system, an automated interactive telephone system, developed collaboratively by the Australian research team and researchers at the Medical Information Systems Unit, Boston University, USA. The Australian TLC Diabetes system has been designed to improve diabetes management by targeting the following key self-management behaviours: blood glucose testing, nutrition, physical activity and medication-taking. Users were asked to call the system weekly using a landline or mobile phone. TLC's responses, including feedback and encouragement, were tailored according to information entered in the TLC database at the start and the answers that it received from participants during all calls.

Table 1 Inclusion and exclusion criteria for study recruitment

Inclusion criteria	Exclusion criteria
Eligibility Step 1	
Type 2 diabetes diagnosis of ≥ 3 months	Diagnosed with dementia/psychiatric co-morbidity
Aged 18–70 years	Currently enrolled in another intervention trial
Residing in greater Brisbane area, Australia	Undergone bariatric surgery in past 2 years
Stable diabetes pharmacotherapy type for ≥ 3 months	Pregnant, lactating, or planning to become pregnant within next 12 months
Ability to clearly speak/understand English via telephone	Diagnosed with condition likely to be fatal within 1 year
Stable pharmacotherapy dosage for ≥ 4 weeks	
Weekly access to telephone	
Eligibility Step 2	
HbA_{1c} ≥ 7.5%	

Training to use the TLC system

The TLC Coordinator met with participants within one week of their baseline data collection to instruct them on the use of the TLC Diabetes kit (containing the TLC Handbook, an ACCU-CHEK® Advantage glucose meter, test strips, and a Bluetooth™ device for uploading their blood glucose results to the TLC Diabetes system). For current smokers, a smoking cessation information pack was also provided. During this session, participants completed a training call to the TLC Diabetes system. Participants were asked to conduct all blood glucose self-monitoring with the study glucose meter and to upload its readings immediately preceding their weekly telephone conversations with the TLC system. Each participant chose a unique personal password that they keyed in at the start of each call that linked the call to their database file and ensured correct participant identification and confidentiality. Before the participants' first call to the TLC system, the TLC Coordinator obtained self-care clinical targets for the participants from their primary healthcare provider (including recommended number of weekly blood glucose tests and blood glucose range, and clearance for physical activity).

Content of weekly telephone calls

Participants were requested to make weekly calls to the system over six months, with calls lasting five-20 minutes, depending upon the call content and participant responses. Blood glucose monitoring was the first topic covered in each weekly call. It was followed by one of three other topics, with these being medication-taking, physical activity or healthy eating (calls 9 to 12 and 21 to 24). When diabetes medication was not prescribed, the medication-taking topic was replaced with physical activity. When clearance for physical activity was not provided by the patient's treating physician, physical activity was replaced by medication-taking. In cases when there was no clearance for physical activity and no pharmaceutical treatment of diabetes, the participant did not hear a second topic on some calls.

TLC Coordinator

The TLC Coordinator briefly telephoned participants after the first two calls and at weeks six, 12, and 20, to identify and resolve any technical issues with the TLC Diabetes system or to determine reasons for not calling. In addition, the TLC Diabetes system sent email "alerts" to a dedicated study email address if any unusual clinical or other issues arose during the conversations, for example, where there were two or more hypoglycaemic levels in the past week. In this instance, the Coordinator would advise the participant of the importance of visiting their primary care physician. More detail on the intervention is available elsewhere [27].

Measurement

Participants in both arms completed comprehensive clinical and self-report assessments at baseline (Time 1), six months following baseline (Time 2), and at 12 months (Time 3); this paper presents the baseline characteristics and six-month primary outcome findings.

Outcome variables

The primary outcomes were HbA_{1c} measured by fasting blood tests taken at the hospital appointment, and HRQL assessed by the participants' self-completion of the SF-36 version 2 (divided into mental and physical component summary scores) [28].

Figure 1 illustrates the stages of recruitment and randomisation.

Representativeness of study sample (Table 2)

To examine the representativeness of the Australian TLC Diabetes sample, the baseline characteristics were compared with data from the Australian Diabetes, Obesity and Lifestyle (AusDiab) study [29], the largest national, population-based sample of Australians measuring the

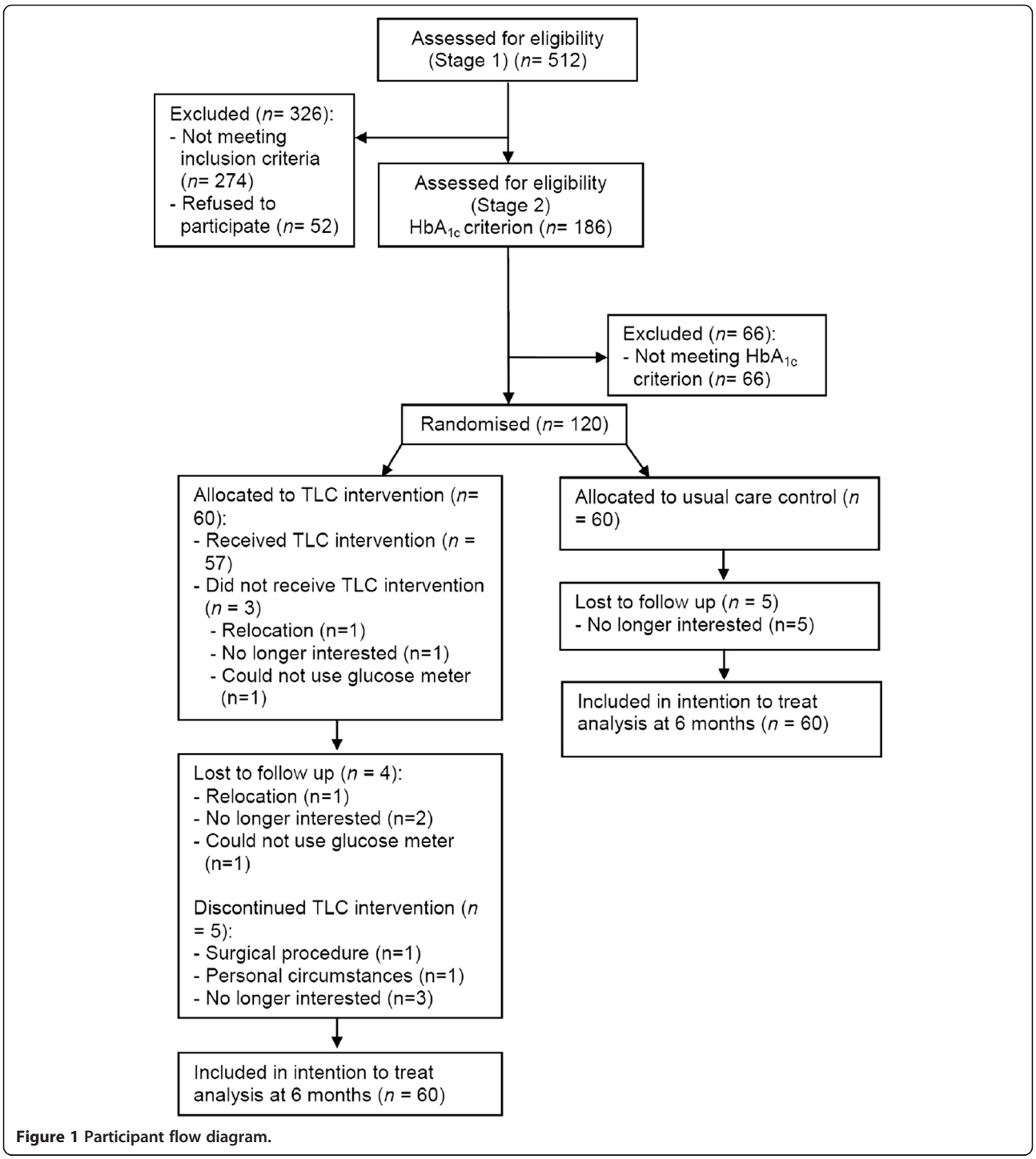


Figure 1 Participant flow diagram.

overall prevalence of diabetes and other chronic conditions. The AusDiab baseline study was conducted during 1999–2000 with data from 11,247 adults [29]. Demographic and behavioural data were collected during interview, and diabetes status was assessed using fasting plasma glucose and oral glucose tolerance tests. A subsample of this nationally representative study, those identified as having diabetes (and based on TLC inclusion criteria), provides the best comparison for the TLC study sample.

Statistical analyses

Detailed power calculations were described in an earlier paper [27], indicating the need to recruit 340 participants to detect a small clinical change of 0.4% in HbA_{1c} with 90% power assuming a 30% rate of loss to

Table 2 Baseline characteristics of Australian Telephone-Linked Care (TLC) Diabetes sample

		TLC Diabetes Intervention (n = 60)	Usual care (n = 60)	Total TLC sample (n = 120)	AusDiab sub-sample (n = 156)
Demographic variables					
Age		58.4 (8.2)	56.4 (8.3)	57.4 (8.3)	56.6 (8.8)
Sex	% male	61.7	63.3	62.5	59.8
Country of birth	% born in Australia	71.7	68.3	70.0	66.7
Marital status	% cohabiting	75.4	74.6	75.0	67.9
Employment	% working	46.7	45.0	45.8	47.7
income	% > $40,000	46.7	51.7	49.2	
Education	% > secondary school	60	70	65.0	55.8
Private medical insurance	% with	56.7	55.0	55.8	47.3
Psychosocial risk factors					
Depression	Low	66.7	78.3	72.5	-
	Intermediate	26.7	20.0	23.3	
	High	6.7	1.7	4.2	
Anxiety	Low	90.0	88.3	89.2	-
	Intermediate	6.7	8.3	7.5	
	High	3.3	3.3	3.3	
Social support	% low	20.0	21.7	20.8	NC
Nutritional self-efficacy		15.2 (3.0)	15.0 (3.2)	15.1 (3.1)	NC
Physical activity self-efficacy		12.8 (3.3)	12.7 (3.5)	12.7 (3.4)	NC
HRQL	Physical component summary	43.7 (8.4)	43.8 (10.2)	43.6 (9.3)	45.2 (12.7)
	Mental component summary	49.8 (8.7)	49.5 (9.1)	49.6 (8.9)	49.5 (9.8)
Health behaviours					
Smoking status	% never	51.7	53.3	52.5	34.0
	% ex-smoker	45.0	46.7	45.0	45.5
	% current	3.3	0	1.7	17.9*
Physical activity	% none	5.0	5.2	5.1	22.4
	% do not meet guidelines	35.0	43.1	39.0	31.4
	% meet guidelines	60.0	51.7	55.9	46.2*
Diet (n = 110)	Energy (kJ/day) ‡	7658 (5884–9745)	7811 (6080–9566)	7704 (6025–9638)	7467 (5850–9455)
	Fibre (g/day) ‡	23 (18–29)	23 (17–29)	23 (17–29)	23 (17–30)
	Fat (g/day) ‡	73 (53–93)	76 (63–95)	75 (57–94)	71 (56–94)
	Saturated fat (g/day) ‡	27 (21–37)	30 (23–38)	29 (22–38)	28 (22–38)
Self-care					
	% adherence to blood glucose testing				
	recommendations	40.0	30.0	36.2	NC
	% checked feet everyday	31.7	20.0	26.1	NC
	% insulin/diabetes medical adherence everyday	87.9	86.0	84.0	NC
Self-reported health	% ≥ good	74.9	73.3	74.2	65.8
Medication use					
Inject insulin	% on insulin	41.7	45.0	43.3	NC

Table 2 Baseline characteristics of Australian Telephone-Linked Care (TLC) Diabetes sample *(Continued)*

Clinical measures				
Systolic blood pressure (mmHg)	135.4 (15.0)	137.0 (15.0)	136.2 (15.0)	140.2 (18.9)*
Diastolic blood pressure (mmHg)	32.5 (28.7-35.9)	32.9 (29.2-37.8)	33.6 (28.8-36.9)	30.0 (26.5-34.6)*
Body mass index (kg/m^2) ‡	107.7 (100.0-114.6)	113.1 (101.7-122.4)	111.0 (101.5-118.7)	103.2 (92.0-115.2)*
Waist circumference (cm) ‡	8.6 (8.0-9.2)	8.5 (7.9-9.5)	8.5 (7.9-9.3)	8.8 (8.1-9.8)*
Glycosylated haemoglobin (HbA_{1c}) (%)‡	9.8 (8.4-11.1)	9.5 (8.1-12.0)	9.6 (8.2-11.4)	11.5 (9.8-14.1)*
Fasting glucose (mmol/l) ‡	15.0 (9.6-24.0)	13.0 (8.3-22.8)	14.0 (9.1-23.8)	NC
Fasting insulin (mU/l) ‡	2.3 (1.4-3.6)	1.9 (1.2-3.0)	2.2 (1.3-3.3)	NC
HOMA Insulin Resistance ‡	4.0 (3.5-4.9)	4.0 (3.4-5.2)	4.0 (3.5-5.2)	5.6 (4.8-6.3) *
Total cholesterol (mmol/l) ‡	32.5 (28.7-35.9)	32.9 (29.2-37.8)	33.6 (28.8-36.9)	30.0 (26.5-34.6)*
High density lipoprotein cholesterol (mmol/l) ‡	1.0 (0.8-1.1)	1.0 (0.8-1.2)	1.0 (0.8-1.1)	1.2 (1.0-1.4) *
Low density lipoprotein cholesterol (mmol/l) ‡	2.1 (1.8-3.1)	2.2 (1.6-3.0)	2.1 (1.7-3.0)	3.2 (2.7-3.9)*
Triglycerides (mmol/l) ‡	1.6 (1–2.1)	1.5 (1.1-2.0)	1.5 (1.1-2.1)	2.0 (1.3-2.9) *
Creatinine (μmol) ‡	83.0 (64.8-98.5)	73.0 (62.0-86.8)	78.0 (62.8-95.3)	83.5 (73.0-92.8)*
Estimated glomerular filtration rate (ml/min) ‡ †	76.0 (64.0-91.0)	85.5 (77.3-91.0)*	83.0 (70.8-91.0)	78.1 (69.5-87.7)*
Clinical history (self-report)				
Doctor-diagnosed hypertension (%)	63.3	68.3	65.8	46.8
Doctor-diagnosed hypercholesterolaemia (%)	60.0	66.7	63.3	48.0
Doctor-diagnosed diabetic eye complications (%)	15.0	21.7	18.3	NC
Doctor-diagnosed diabetic neuropathy (%)	18.3	25.0	21.7	NC
Doctor-diagnosed kidney disease (%)	11.7	5.0	8.3	NC
Doctor-diagnosed cardiovascular disease (%)	28.3	30.0	29.2	NC

Data are presented as means (SD) and percentages, or as ‡medians (inter-quartile range) for skewed data. Group comparisons between TLC study arms and between TLC and AusDiab samples of normally distributed data used independent samples t-tests and chi square tests. Group comparisons between TLC study arms and between TLC and AusDiab samples of non-normally distributed variables used Mann–Whitney U test, *p <0.05.

† Estimated glomerular filtration rate data highly skewed (values over 90 ml/min labelled 91).

HRQL: Health-related quality of life.

HOMA: Homeostasis model assessment.

High risk AusDiab group inclusion criteria are type 2 diabetes, within TLC age-range, and $HbA_{1c} \geq 7.5$ %; TLC-AusDiab group comparison are made with full TLC sample (n = 120).

NC: Not comparable - missing comparisons with AusDiab subsample due to incomparable methods of data collection between studies.

follow-up. However, due to the slowness of recruitment (described below), our final sample comprised a total of 120 participants (60 per study arm). Although the comprehensive recruitment effort achieved a very good response from individuals with diabetes (n = 512), a large proportion of these respondents either did not wish to participate or were ineligible due to either not meeting the HbA_{1c} or age eligibility criteria. Recruitment was stopped after 18 months with 120 participants having been recruited. Therefore, the power calculations were re-evaluated based on this number of participants, again assuming 30% loss to follow-up. With 80% power and a type 1 error of 5% (two-tailed), a difference in our primary outcome, HbA_{1c}, of 0.61% between the intervention and control arms (based on a standard deviation change of 1.0% between the randomised arms) at 12-month follow-up can be detected. This effect size would indicate a feasible outcome of clinical significance [30] for the intervention.

For the analysis of the six-month results, HbA_{1c} values were logarithmically transformed in order to achieve an approximate normal distribution. Analyses of covariance were used to examine the effects of the intervention (study arm allocation) on the primary outcomes (log HbA_{1c} and HRQL), with the inclusion of baseline values of the outcomes as covariates. Results for HbA_{1c} are presented as geometric means for each study arm and as a ratio of geometric means when comparing study arms. The geometric mean is a natural quantity to use for presenting the centre of skewed data and is computed by exponentiating the average of the logarithmically transformed HbA_{1c} values [31]. To assess heterogeneity of the effect of TLC according to baseline values, interactions between study arm allocation and baseline values were included in further regression models. Creatinine and e-GFR were included as covariates in these analyses, since their levels at baseline differed sizeably between study arms. The sensitivity of conclusions to imbalances in baseline characteristics was assessed via additional ANCOVA analyses adjusting for all characteristics exhibiting any potentially important imbalances. To account for subjects lost to follow-up in intention-to-treat

analyses, multiple imputation was performed using ten imputed datasets [32].

For the comparison of the baseline TLC sample characteristics with the AusDiab study sample, as well as for the attrition comparisons, independent samples t-tests (continuous data) and chi-square tests (categorical data) were used where the data were normally distributed, and Mann–Whitney U tests were employed for highly skewed data. All analyses were performed using SPSS 18.0, with the statistical significance level set at $p < 0.05$.

Results

Of the 52 individuals who did not wish to participate at the initial eligibility screening stage, the primary reason for non-participation was lack of interest (n = 21), with an additional 11 reporting potential difficulties with travel for the baseline data collection. Other reasons included lack of time due to work and other commitments. There were no age differences between those who were willing and unwilling to participate, although there was a higher proportion of women who were unwilling to participate compared with those who chose to participate (61.5% compared with 43%).

As shown in Table 2, which summarises the baseline characteristics of the TLC and usual care arms, the Australian TLC Diabetes sample had a mean age of 57.4 years (± 8.3), with a higher proportion of men (62.5%) than women. The vast majority of participants were born in Australia (70.0%), were married or cohabiting with a partner (75.0%), with education above secondary school level (65.0%). Approximately half of the sample were employed (45.8%) and had complementary private medical insurance (55.8%). The mean number of hours per week spent exercising was reported to be 6.1 (± 6.4), with the majority of the sample (55.9%) participating in the nationally-recommended level of weekly physical activity (>150 minutes of exercise per week in at least 5 sessions per week [33]). Only 1.7% of the sample were current smokers. Approximately three quarters of the sample rated their health as good or higher (74.2%). Nearly two-thirds of the sample had been previously diagnosed by a doctor with hypertension (65.8%) and hypercholesterolaemia (63.3%), and therefore were likely to be receiving treatment for these conditions as was reflected in their blood pressure and lipid profiles that predominantly fell within the normal range.

Comparison of baseline sample characteristics between study arms

The baseline sample characteristics were compared across the usual care and intervention arms to evaluate the randomisation process (Table 2). Comparison of the baseline characteristics across usual care and intervention arms revealed important differences in e-GFR, which showed a significantly greater impairment in renal function in the intervention compared with usual care arm, and creatinine. Other differences observed were in age, education, and self-care behaviours (adherence to blood glucose testing recommendations and daily insulin/diabetes medications, and foot inspections). Adjustments were made for these variables in sensitivity analyses.

Post-intervention results at six months

Attrition

Of the total sample, 92.5% completed the six-month assessment (see Figure 1). Overall, nine participants (two women and seven men) withdrew from participation in the study, four in the intervention arm and five in the usual care arm. The reasons given for withdrawal from the usual care arm were all related to frustration at 'missing out' on the intervention. The participants receiving the Australian TLC Diabetes intervention withdrew for a range of reasons, including relocation, being unable to use the blood glucose meter, and disappointment with the intervention. The sociodemographic, behavioural or biological profiles were compared between those people who remained in the study and the nine people who withdrew. There were no significant differences at baseline across any of the domains of risk factor profiles.

Use of Australian TLC Diabetes system

The mean number of completed calls for the Australian TLC Diabetes participants during the six-month intervention was 18 (± 6), ranging between 2 and 27 calls, with a mean call duration of 11 minutes (± 1). The mean percentage of completed calls out of the expected weekly calls for all individuals in the intervention condition was 76% (± 22). More detailed analyses of the usage of the Australian TLC Diabetes system are beyond the scope of this paper and are to be presented in a future manuscript.

A small number of people in the intervention arm (n = 5) discontinued participation in the intervention but still completed the six-month assessment (Figure 1). Out of these, two participants made less than five calls and one made only seven calls.

Study outcomes

These analyses were based on intention-to-treat. There was a statistically significant difference in HbA_{1c} at six months between the usual care and TLC Diabetes arms. The geometric mean (arithmetic means provided in parentheses) of HbA_{1c} decreased from 8.7% (8.8%) to 7.9% (8.0%) in the TLC Diabetes arm, compared with 8.9% (9.0%) to 8.7% (8.9%) in the usual care arm, with the adjusted ratio of six-month geometric means of 0.91 (95% CI 0.86-0.93, p = 0.002) (Table 3). The ratio of

Table 3 Baseline and post-intervention primary outcome values between usual care and Australian TLC Diabetes arms

	Baseline n = 60	Post-intervention n = 60	Difference between groups* (95% CI, *p*)
HbA_{1c} (%)			Ratio
Usual care	8.9 (8.6-9.2)	8.7 (8.7-9.0)	0.91 (0.86-0.93, $p = 0.002$)
TLC Diabetes	8.7 (8.4-9.0)	7.9 (7.6-8.3)	
Health-related quality of life - mental			
Usual care	49.5 (47.1-50.3)	48.7 (47.1-50.3)	3.0 (0.8-5.2 $p = 0.007$)
TLC Diabetes	49.8 (47.5-52.0)	51.7 (50.2-53.3)	
Health-related quality of life - physical			
Usual care	45.4 (43.0-47.9)	45.2 (43.8-46.6)	0.4 (−1.7-2.4, $p = 0.7$)
TLC Diabetes	45.5 (43.0-47.9)	45.6 (44.1-47.0)	

Data presented in the first two columns are geometric means (95 % CI) for HbA_{1c} values and arithmetic means (95 % CI) for HRQL values. The post-intervention values are adjusted for baseline values, e-GFR and creatinine.
*The result in the last column for HbA_{1c} is the ratio of the geometric means in the TLC Diabetes arm compared with usual care arm.
For HRQL, it is the difference in arithmetic means. All analyses were conducted based on the intention-to-treat principle and adjust for the baseline of the outcome variable, e-GFR and creatinine values.

0.91 means that the geometric mean HbA_{1c} at six months in the TLC arm is 0.91 of the value in the usual care arm after adjustment for baseline covariates. There was slight evidence that the difference in HbA_{1c} at six months between study arms increased with baseline HbA_{1c} ($p = 0.09$ for the interaction term in regression model). This suggested that the difference in six-month HbA_{1c} between TLC and usual care patients was greater in patients with high baseline HbA_{1c} values than in patients with low values. Of participants in the intervention arm, 20 % achieved HbA_{1c} levels of 7.0% or lower (95% CI 9.6-29.7), compared with 15% (95% CI 4.4-24.7) in the usual care arm ($p = 0.32$).

In terms of HRQL, the mental component summary score was found to be significantly different between the two arms at six months (difference = 3.0, $p = 0.007$), after controlling for baseline mental HRQL, plus other covariates (Table 3). Mental HRQL improved in the TLC Diabetes group, compared with those in the usual care group where mental HRQL decreased marginally. There was no interaction between study arm allocation and baseline levels for mental HRQL ($p = 0.4$). No differences were observed in physical HRQL between the usual care and intervention arms ($p = 0.7$).

Comparison of sample characteristics between Australian TLC and AusDiab samples

To determine the representativeness of the TLC sample at baseline, we used a comparable subsample of individuals from the AusDiab study, obtained from applying the Australian TLC Diabetes criteria for age range and HbA_{1c} levels (≥ 7.5%) to the subsample (n = 643) of those classified in AusDiab as having diabetes. 156 AusDiab participants were identified for comparison with the Australian TLC Diabetes sample. Overall, the AusDiab and TLC samples were similar (Table 2). There were no significant differences between the TLC sample and the AusDiab subsample across demographic variables, HRQL, and self-reported health variables. Behaviourally, there were no differences in nutrition self-reports between the study populations, however the TLC sample reported markedly lower smoking rates and were more likely to perform the recommended levels of exercise. In terms of their clinical profiles, the TLC sample appeared healthier, with lower systolic blood pressure, and generally better glucose and lipid profiles. These results, however, are likely to reflect the increased levels of doctor-diagnosed hypertension and hypercholesterolaemia, and therefore probably high levels of treatment in the TLC sample. Interestingly, despite their reported healthier behavioural profiles, the TLC sample were significantly more likely to be obese using both BMI and waist circumference classifications.

Discussion

This randomised controlled trial evaluated the efficacy of an automated, interactive telephone intervention for improving the management of diabetes. As far as we are aware, this is one of the first studies in the world to formally evaluate an automated telephone system for diabetes management that involves tailoring to individual needs and the findings offer promising results for the longer term use of this kind of program for people with diabetes. We have demonstrated that the Australian TLC Diabetes program significantly improved glycaemic control and mental HRQL after six months for those who participated in the program compared with the routine care condition.

Participation in the Australian TLC Diabetes intervention led to a significant improvement of HbA_{1c}, compared with the routine care available to people with diabetes in Brisbane, Australia. The mean reduction in HbA_{1c} of 0.8 % in the intervention arm is of substantial clinical significance if maintained long-term. Results

from the UKPDS study highlight the substantial reductions in all diabetes endpoints associated with 1% reduction in HbA_{1c} [7], such as 21% of deaths related to diabetes, 14% of myocardial infarction and 37% microvascular complications [30]. A meta-analysis reported comparable levels of HbA_{1c} improvement from the pooled effects of 31 previous interventions providing education on self-management of diabetes [9]. The majority of studies cited in the review, however, directly involved healthcare professionals/health workers for the provision of diabetes management education. Another meta-analysis evaluating the use of mobile phone interventions to improve glycaemic control showed a pooled change of 0.5% over six months, however, again with heavy involvement of healthcare personnel for intervention delivery [11]. One previous study of another fully-automated telephone intervention aimed at improving glycaemic control failed to show significant post-intervention differences between intervention and control groups in levels of HbA_{1c} [13]; however, that system did not provide tailored feedback to individuals. Therefore, a major advantage of the Australian TLC Diabetes program is its successful impact on glycaemic control and the potential for reduced costs and increased accessibility associated with an automated telephone-linked system for the provision of tailored diabetes management.

In addition to the observed improvements in glycaemic control, mental HRQL was significantly enhanced in people who received the intervention compared with those who did not, despite this not being a specific focus of the TLC program for the trial. The burden of daily management of diabetes and the development of complications lead to compromised HRQL in populations with diabetes [34,35], and therefore enhancing well-being, in addition to diabetes management per se, is an additionally important outcome. Despite this improvement reflecting only a small effect size (0.20) [36], the literature in this field indicates that even small effect sizes of HRQL improvement may be of clinical significance in the longer term [37-39]. Interestingly, the physical component of HRQL did not improve during the six-month intervention period. A brief computer-assisted diabetes self-management intervention on quality of life outcomes showed no change in HRQL, however, their two-month follow-up might not have been long enough to detect changes [40]. In contrast, the pooled results from 20 publications showed that people with diabetes experience improved HRQL after receiving interventions designed to develop their diabetes self-management behaviours [37], although this meta-analysis did not differentiate between the mental and physical components of HRQL.

Another important aspect of this study is the focus on people with poor glycaemic control ($HbA_{1c} \geq 7.5\%$), indicating difficulty in their self-management of diabetes with the available routine care. These people are likely to be most at risk of the development of complications associated with diabetes, and therefore, given the results achieved, Australian TLC Diabetes has the potential to improve the health of the highest risk groups. Consequently, this program also provides the opportunity to significantly reduce the financial burden of type 2 diabetes on the healthcare system. Subsequent analyses will examine the cost-effectiveness of the program, which will have important implications for the widespread implementation of the program.

Our comparison of the TLC sample with a 'matched' subgroup from the AusDiab study sample suggests that the TLC participants did not differ significantly in terms of demographic characteristics from the best available data from a general population-based diabetes sample in Australia. The baseline AusDiab study, conducted in 1999–2000, offers benchmark national data on the prevalence of diabetes, obesity, hypertension, and kidney disease in Australia. This indicates the representativeness and external validity of our results and their applicability to other diabetes populations.

The trial was completed in accordance with the Medical Research Council's guidelines for the effective design and evaluation of complex intervention trials [41]. Principal components of any effective complex intervention include feasibility, participant-engagement, identification of mechanisms for intervention outcomes, and trial fidelity [42]. The feasibility and relevance of the Australian TLC Diabetes program are demonstrable within the current context of type 2 diabetes. The accessibility of the telephone-delivered intervention over the long-term is particularly important for a widespread chronic condition, such as diabetes, which requires ongoing management and affects a large proportion of the population. The very high usage of the Australian TLC Diabetes system and results to date indicate that the participants in the intervention arm engaged with the program, with over three quarters of weekly calls being completed. Full details of system usage were recorded as part of the data collection and will be reported elsewhere for full process evaluation of the system's usability and participant satisfaction, as well as whether the cost of the intervention provides acceptable value for money. Furthermore, the intervention was able to affect pathways that led to improvements in glycosylated haemoglobin and therefore diabetes management, as well as improvement in mental health-related quality of life for the participants. The fidelity of the trial implementation in accordance with the original design and protocol [27] was strong. Difficulties were encountered during recruitment and this led to increased recruitment opportunities via enhanced presence at Diabetes Australia – Queensland shops and seminars and hospital

diabetes clinics. The sample size was smaller than originally planned, however, as discussed, the sample obtained is powered to detect group differences that will be both statistically and clinically significant at 12-month follow-up. No changes were applied regarding the randomisation process or implementation of the intervention.

Although only glomerular filtration rate significantly varied across the study arms at baseline, other baseline characteristics (Table 2) showed some differences. Separate analyses tested the impact of the inclusion of these variables individually on the main results and the main outcome results did not change. As with most research, it is possible that a selection bias operated in this study, with people willing to participate being more likely to prioritise their health and/or have the social, educational, and economic resources to accommodate participation. The study requirement of access to a telephone meant that there may have been a socioeconomic selection bias; however in the geographic area from which we recruited, over 96% of households have a fixed phone connection, so we are confident that this criterion did not appreciably influence participation. It is also possible that the reduced sample size and some of the challenges associated with trial recruitment may limit generalisability. More research is required to investigate generalisability and to explore uptake by others with diabetes. Although there was a suggestion of an increasing effect of intervention with increasing baseline HbA_{1c} values (from the interaction test), this did not reach conventional levels of statistical significance and should be reassessed in future studies.

A substantial body of research conducted over the last 30 years has drawn attention to the importance of ongoing support and follow-up to sustain improvements in diabetes management and management of other chronic conditions, with strong links to health and self-care behaviours [43-45]. Therefore a diabetes management support program such as this, designed to provide easy access to long-term (potentially cost-effective) support, is of paramount importance, and hence, this kind of program also requires detailed evaluation in the longer term as well. A subsequent paper will elucidate the changes in behaviour that may have facilitated the improvements observed.

Conclusions

Our results indicate that the six-month Australian TLC Diabetes program led to improvements in diabetes management, with significant benefits to mental health functioning and improved glycaemic control. If these results were maintained long term, such results would be expected to lead to important reductions in diabetes-related complications and mortality [30]. With the increasing accessibility to and feasibility of such telehealth interventions, the TLC program has excellent potential to be 'scaled up' and deliverable to large numbers of individuals with diabetes.

Competing interests

Dr. Friedman has stock ownership and a consulting agreement with Infomedics, the company that owns commercial rights to the TLC technology used in the computerized intervention. He is also a member of its Board of Directors. The other authors declare that they have no competing interests.

Authors' contributions

EDW analysed the data and wrote the manuscript. DB collected the data, contributed to study development, discussion and manuscript writing. AF, AR, SA, PS, RF and BO contributed to study development, and discussion, reviewing/editing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The study is funded by a National Health Medical Research Council project grant (ID 443214), by the HCF Health and Medical Research Foundation and by Queensland Health. We wish to thank all study participants, Diabetes Australia for provision of educational material and Diabetes Australia - Queensland for its assistance with recruitment. We acknowledge the dedication of the investigator team: Prof Mary Courtney, Prof Richard Wootton, and Prof Kerrie Mengersen. We also acknowledge the commitment of the project staff: Mandy Cassimatis, Lyndall Kopp, Megan Rollo, Wei-I Wu, Adrienne O'Neil and Vivien Harris. We would like to acknowledge the Baker IDI Heart and Diabetes Institute for access to the AusDiab data. Furthermore, we would like to thank Dr Stephan Gaedhe, one of the principal developers of the TLC Diabetes system, and the Australian Diabetes Educators Association for their contributions. We are grateful for the input from staff at the Medical Information Systems Unit, Boston University. Finally, we also wish to thank Roche Diagnostics ACCUCHEK for their supply of the glucose meters and Alive Technologies for all of their technical advice. EDW was funded by a Diabetes UK Moffat Fellowship.

Author details

[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia. [2]School of Medicine, University of Queensland, Brisbane, Australia. [3]Department of Diabetes and Endocrinology, Princess Alexandra Hospital, Brisbane, Australia. [4]Institute of Health and Biomedical Innovation, Queensland University of Technology, Queensland, Australia. [5]School of Exercise and Nutrition Sciences, Queensland University of Technology, Brisbane, Australia. [6]Centre for Applied Health Economics, School of Medicine, Griffith Health Institute, Griffith University, Queensland, Australia.

Received: 11 November 2011 Accepted: 22 June 2012
Published: 3 August 2012

References

1. Shaw JE, Sicree RA, Zimmet PZ: **Global estimates of the prevalence of diabetes for 2010 and 2030.** *Diabetes Res Clin Pract* 2010, **87**:4–14.
2. Zhang P, Zhang X, Brown J, Vistisen D, Sicree R, Shaw J, Nichols G: **Global healthcare expenditure on diabetes for 2010 and 2030.** *Diabetes Res Clin Pract* 2010, **87**:293–301.
3. American Diabetes Association: **Standards of Medical Care in Diabetes –2011.** *Diabetes Care* 2011, **34**:S11–S61.
4. Kemp TM, Barr EL, Zimmet PZ, Cameron AJ, Welborn TA, Colagiuri S, Phillips P, Shaw JE: **Glucose, lipid, and blood pressure control in Australian adults with type 2 diabetes: the 1999–2000 AusDiab.** *Diabetes Care* 2005, **28**:1490–1492.
5. Saydah SH, Fradkin J, Cowie CC: **Poor control of risk factors for vascular disease among adults with previously diagnosed diabetes.** *JAMA* 2004, **291**:335–342.
6. Selvin E, Marinopoulos S, Berkenblit G, Rami T, Brancati FL, Powe NR, Golden SH: **Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus.** *Ann Intern Med* 2004, **141**:421–431.

7. UKPDS G: **Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33).** *Lancet* 1998, **352**:837–853.
8. Funnell MM, Brown TL, Childs BP, Haas LB, Hosey GM, Jensen B, Maryniuk M, Peyrot M, Piette JD, Reader D, *et al*: **National Standards for Diabetes Self-Management Education.** *Diabetes Care* 2009, **33**:S89–S96.
9. Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM: **Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control.** *Diabetes Care* 2002, **25**:1159–1171.
10. Eakin EG, Lawler SP, Vandelanotte C, Owen N: **Telephone interventions for physical activity and dietary behavior change: a systematic review.** *Am J Prev Med* 2007, **32**:419–434.
11. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, Huang J, Wang L, Gu D: **Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis.** *Diabetic Med* 2010, **28**:455–463.
12. Murray E, Burns J, See TS, Lai R, Nazareth I: **Interactive Health Communication Applications for people with chronic disease.** *Cochrane Database Syst Rev* 2005, **4**:CD004274.
13. Graziano JA, Gross CR: **A randomized controlled trial of an automated telephone intervention to improve glycemic control in type 2 diabetes.** *ANS Adv Nurs Sci* 2009, **32**:E42–E57.
14. Piette JD, Weinberger M, Kraemer FB, McPhee SJ: **Impact of automated calls with nurse follow-up on diabetes treatment outcomes in a Department of Veterans Affairs Health Care System: a randomized controlled trial.** *Diabetes Care* 2001, **24**:202–208.
15. Shea S, Weinstock RS, Teresi JA, Palmas W, Starren J, Cimino JJ, Lai AM, Field L, Morin PC, Goland R, *et al*: **A Randomized Trial Comparing Telemedicine Case Management with Usual Care in Older, Ethnically Diverse, Medically Underserved Patients with Diabetes Mellitus: 5 Year Results of the IDEATel Study.** *J Am Med Inform Assoc* 2009, **16**:446–456.
16. Friedman RH, Stollerman JE, Mahoney DM, Rozenblyum L: **The virtual visit: using telecommunications technology to take care of patients.** *J Am Med Inform Assoc* 1997, **4**:413–425.
17. Farzanfar R, Frishkopf S, Migneault J, Friedman R: **Telephone-linked care for physical activity: a qualitative evaluation of the use patterns of an information technology program for patients.** *J Biomed Inform* 2005, **38**:220–228.
18. Rubin A, Migneault JP, Marks L, Goldstein E, Ludena K, Friedman RH: **Automated telephone screening for problem drinking.** *J Stud Alcohol* 2006, **67**:454–457.
19. Delichatsios HK, Friedman RH, Glanz K, Tennstedt S, Smigelski C, Pinto BM, Kelley H, Gillman MW: **Randomized trial of a "talking computer" to improve adults' eating habits.** *Am J Health Promot* 2001, **15**:215–224.
20. Jarvis KL, Friedman RH, Heeren T, Cullinane PM: **Older women and physical activity: using the telephone to walk.** *Womens Health Iss* 1997, **7**:24–29.
21. King AC, Friedman R, Marcus B, Castro C, Napolitano M, Ahn D, Baker L: **Ongoing physical activity advice by humans versus computers: the Community Health Advice by Telephone (CHAT) trial.** *Health Psychol* 2007, **26**:718–727.
22. Pinto BM, Friedman R, Marcus BH, Kelley H, Tennstedt S, Gillman MW: **Effects of a computer-based, telephone-counseling system on physical activity.** *Am J Prev Med* 2002, **23**:113–120.
23. DeMolles DA, Sparrow D, Gottlieb DJ, Friedman R: **A pilot trial of a telecommunications system in sleep apnea management.** *Med Care* 2004, **42**:764–769.
24. Mahoney D, Tennstedt S, Friedman R, Heeren T: **An automated telephone system for monitoring the functional status of community-residing elders.** *Gerontologist* 1999, **39**:229–234.
25. Friedman RH, Kazis LE, Jette A, Smith MB, Stollerman J, Torgerson J, Carey K: **A telecommunications system for monitoring and counseling patients with hypertension. Impact on medication adherence and blood pressure control.** *Am J Hypertens* 1996, **9**:285–292.
26. Houlihan BV, Jette A, Paasche-Orlow M, Wierbicky J, Ducharme S, Zazula J, Cuevas P, Friedman RH, Williams S: **A Telerehabilitation Intervention for Persons with Spinal Cord Dysfunction.** *Am J Phys Med Rehab* 2011, doi:10.1097/PHM.1090b1013e31820b31140f. Publish Ahead of Print.
27. Bird D, Oldenburg B, Cassimatis M, Russell A, Ash S, Courtney MD, Scuffham PA, Stewart I, Wootton R, Friedman RH: **Randomised controlled trial of an automated, interactive telephone intervention to improve type 2 diabetes self-management (Telephone-Linked Care Diabetes Project): study protocol.** *BMC Publ Health* 2010, **10**:599.
28. Ware JE Jr: **Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.** *Med Care* 1992, **30**:473–483.
29. Dunstan DW, Zimmet PZ, Welborn TA, Cameron AJ, Shaw J, de Courten M, Jolley D, McCarty DJ: **The Australian Diabetes, Obesity and Lifestyle Study (AusDiab)–methods and response rates.** *Diabetes Res Clin Pract* 2002, **57**:119–129.
30. Stratton IM, Adler AI, Neil HAW, Matthews DR, Manley SE, Cull CA, Hadden D, Turner RC, Holman RR: **Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study.** *BMJ* 2000, **321**:405–412.
31. Altman D: *Practical Statistics for Medical Research*. London: Chapman and Hall; 1990.
32. Little RJ, Rubin DB: *Statistical analysis with missing data*. 2nd edition. New York: Wiley; 2002.
33. Armstrong T: *Bauman A*. Davies J: Physical activity patterns of Australian adults. Edited by Australian Institute of Health and Welfare. Canberra; 2000.
34. Rubin RR, Peyrot M: **Quality of life and diabetes.** *Diabetes Metab Res Rev* 1999, **15**:205–218.
35. Wexler DJ, Grant RW, Wittenberg E, Bosch JL, Cagliero E, Delahanty L, Blais MA, Meigs JB: **Correlates of health-related quality of life in type 2 diabetes.** *Diabetologia* 2006, **49**:1489–1497.
36. Cohen J: *Statistical power analysis for the behavioral sciences*. 2nd edition. Hillsdale (NJ): Lawrence Erlbaum Assoc; 1988.
37. Cochran J, Conn VS: **Meta-analysis of Quality of Life Outcomes Following Diabetes Self-management Training.** *The Diabetes Educator* 2008, **34**:815–823.
38. O'Neil A, Sanderson K, Oldenburg B, Taylor CB: **Impact of Depression Treatment on Mental and Physical Health-Related Quality of Life of Cardiac Patients: A meta-analysis.** *J Cardiopulm Rehab Prev* 2011, **31**:146–156.
39. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D: **Determining Clinically Important Differences in Health Status Measures: A General Approach with Illustration to the Health Utilities Index Mark II.** *PharmacoEconomics* 1999, **15**:141–155.
40. Glasgow RE, Nutting PA, Toobert DJ, King DK, Strycker LA, Jex M, O'Neill C, Whitesides H, Merenich J: **Effects of a brief computer-assisted diabetes self-management intervention on dietary, biological and quality-of-life outcomes.** *Chronic Illn* 2006, **2**:27–38.
41. Medical Research Council: *A framework for development and evaluation of RCTs for complex interventions to improve health*. London: Medical Research Council; 2000.
42. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, Guthrie B, Lester H, Wilson P, Kinmonth AL: **Designing and evaluating complex interventions to improve health care.** *BMJ* 2007, **334**:455–459.
43. Clark M: **Diabetes self-management education: a review of published studies.** *Prim Care Diabetes* 2008, **2**:113–120.
44. Minet L, Møller S, Vach W, Wagner L, Henriksen JE: **Mediating the effect of self-care management intervention in type 2 diabetes: A meta-analysis of 47 randomised controlled trials.** *Patient Educ Couns* 2010, **80**:29–41.
45. Norris SL, Engelgau MM, Narayan KM: **Effectiveness of self-management training in type 2 diabetes: a systematic review of randomized controlled trials.** *Diabetes Care* 2001, **24**:561–587.

doi:10.1186/1471-2458-12-602
Cite this article as: Williams *et al.*: Randomised controlled trial of an automated, interactive telephone intervention (TLC Diabetes) to improve type 2 diabetes management: baseline findings and six-month outcomes. *BMC Public Health* 2012 **12**:602.

Risk of Thromboembolism, Recurrent Hemorrhage, and Death After Warfarin Therapy Interruption for Gastrointestinal Tract Bleeding

Daniel M. Witt, PharmD, FCCP, BCPS; Thomas Delate, PhD; David A. Garcia, MD; Nathan P. Clark, PharmD; Elaine M. Hylek, MD; Walter Ageno, MD; Francesco Dentali, MD; Mark A. Crowther, MD

Background: Patients who not only survive a warfarin-associated gastrointestinal tract bleeding (GIB) event but also have an ongoing risk for thromboembolism present 2 clinical dilemmas: whether and when to resume anticoagulation. The objective of this study was to determine the incidence of thrombosis, recurrent GIB, and death, as well as the time to resumption of anticoagulant therapy, during the 90 days following a GIB event.

Methods: In this retrospective, cohort study using administrative and clinical databases, patients experiencing GIB during warfarin therapy were categorized according to whether they resumed warfarin therapy after GIB and followed up for 90 days. Variables describing the management and severity of the index GIB were also collected. Kaplan-Meier curves were constructed to estimate the survival function of thrombosis, recurrent GIB, and death between the "resumed warfarin therapy" and "did not resume warfarin therapy" groups, with Cox proportional hazards modeling to adjust for potentially confounding factors.

Results: There were 442 patients with warfarin-associated index GIB included in the analyses. Following the index GIB, 260 patients (58.8%) resumed warfarin therapy. Warfarin therapy resumption after the index GIB was associated with a lower adjusted risk for thrombosis (hazard ratio [HR], 0.05; 95% CI, 0.01-0.58) and death (HR, 0.31; 95% CI, 0.15-0.62), without significantly increasing the risk for recurrent GIB (HR, 1.32; 95% CI, 0.50-3.57).

Conclusions: The decision to not resume warfarin therapy in the 90 days following a GIB event is associated with increased risk for thrombosis and death. For many patients who have experienced warfarin-associated GIB, the benefits of resuming anticoagulant therapy will outweigh the risks.

Arch Intern Med. 2012;172(19):1484-1491.
Published online September 17, 2012.
doi:10.1001/archinternmed.2012.4261

GASTROINTESTINAL TRACT bleeding (GIB) affects an estimated 4.5% of warfarin-treated patients annually and is associated with a significant risk for death.[1] A history of major bleeding is an important predictor for future serious bleeding, suggesting that patients with GIB might be considered for discontinuation of warfarin therapy.[2,3] However, interruption or permanent discontinuation of warfarin therapy increases the risk of thromboembolic complications.[4] Thus, patients with warfarin-associated GIB present 2 clinical dilemmas: should warfarin therapy be stopped and, if so, when should it be resumed?

Although some have suggested that anticoagulation can be safely restarted relatively soon after a major bleeding event, there is neither high-quality evidence nor consensus about the ideal timing or risk of reanticoagulation. Indeed, surprisingly little is known about warfarin therapy interruption and resumption following GIB.[2,5] An observational study of patients newly initiated on anticoagulation therapy for venous thrombosis who had

See Invited Commentary at end of article

major bleeding reported an association between resuming anticoagulation and rebleeding.[6] However, patients whose index bleeding event occurred during long-term anticoagulation therapy were not evaluated. Other studies examining the resumption of anticoagulation therapy following major bleeding have been limited by small numbers of patients, selection bias, or both.[2,5,7,8]

In this study of warfarin-treated patients who experienced GIB, we sought to

Author Affiliations are listed at the end of this article.

determine the incidence of subsequent thrombosis, recurrent GIB, and death as well as the time to resumption of warfarin therapy during 90 days of follow-up. We evaluated patient characteristics as well as the duration of warfarin therapy interruption to identify factors associated with thromboembolism, recurrent GIB, or death.

METHODS

STUDY DESIGN AND OUTCOMES OF INTEREST

We conducted a retrospective, cohort study using administrative and clinical databases from Kaiser Permanente Colorado (KPCO). Anticoagulation services at KPCO are provided by a centralized Clinical Pharmacy Anticoagulation Service (CPAS).[9] We used integrated, electronic medical, pharmacy, and laboratory record systems along with the CPAS database (DawnAC; 4S Systems Ltd) to identify patients, treatments, and outcomes for this study. Approval to conduct this study was obtained from the KPCO institutional review board.

We used administrative coding data (see the eAppendix for a complete listing of *International Classification of Diseases, Ninth Revision* [*ICD-9*] codes; http://www.archinternmed.com) to identify adult KPCO members who (1) were hospitalized or had an emergency department (ED) encounter for GIB (index GIB) between January 1, 2005, and December 31, 2008; (2) had an outpatient purchase of warfarin and an international normalized ratio (INR) in the 60 days prior to the index GIB; (3) had continuous KPCO membership in the 180 days prior to and 90 days after the index GIB (patients who died within 90 days after the index GIB were included); and (4) did not have a GIB diagnosis recorded during the 6 months prior to the index GIB.

The primary outcomes of interest were thrombosis (stroke, systemic embolism, and venous thromboembolism), recurrent GIB, and death from any cause during the 90 days following index GIB. Gastrointestinal tract bleeding (index and recurrent) and thrombotic events were identified first through electronic queries of inpatient and ED claims databases using *ICD-9* codes (see eAppendix) and then confirmed via manual medical record review by study investigators (D.M.W. and N.P.C.) using a standardized abstraction form. Validation of study outcomes required objective evidence of either clinically overt gastrointestinal tract hemorrhage (eg, visualization of blood in stool, vomit, or gastric aspirate; positive guaiac test result; evidence from endoscopy or colonoscopy) or thrombosis (eg, positive computed tomographic scan, magnetic resonance image, ventilation-perfusion scan, or ultrasonogram). Date and cause of death were ascertained from death certificates and medical record review. All records were independently reviewed by 2 investigators (D.M.W. and N.P.C.), with disagreements resolved by a third reviewer (E.M.H.).

The following variables describing the management and severity of the index GIB were collected: presentation INR, warfarin therapy interruption, plasma or blood transfusion, phytonadione administration, intensive care unit (ICU) admission, length of ED/inpatient stay, and warfarin therapy resumption. We also recorded age, sex, warfarin indication, INR range, time from warfarin therapy initiation to index GIB, aspirin and nonsteroidal anti-inflammatory drug (NSAID) use before index GIB, low-molecular-weight heparin use after index GIB, and proportion of INR values in range during the 3 months before index GIB. Information about comorbidities was collected using *ICD-9* codes. A validated aggregate measure of patient comorbidity, the Chronic Disease Score (CDS), was calculated for each patient using ambulatory prescription medication data recorded before the index GIB.[10,11] For patients with atrial fibrillation, the $CHADS_2$ score, a clinical prediction rule for estimating the risk of stroke, was calculated by assigning 1 point for diagnoses of congestive heart failure, hypertension, age 75 years or older, and diabetes mellitus and 2 points for prior stroke or transient ischemic attack.[12]

STATISTICAL ANALYSES

All patients were assigned to 1 of 2 groups defined by warfarin therapy resumption after the index GIB (ie, "resumed warfarin therapy" and "did not resume warfarin therapy" groups). When warfarin therapy was not interrupted, patients were included in the resumed warfarin therapy group. Categorical data were reported as percentages, and continuous data were reported as means (standard deviations)) and medians (interquartile ranges [IQRs]). Comparisons between groups for categorical data were made with the χ^2 or Fisher exact tests, whereas continuous data were compared using 2-sample *t* tests or Wilcoxon rank sum tests. Kaplan-Meier curves were constructed to estimate the survival function of thrombosis, recurrent GIB, and death between the resumed warfarin therapy and did not resume warfarin therapy groups. Patients were censored at thrombosis, recurrent GIB, death, or 90 days after index GIB, whichever came first.

A propensity score[13] for resumption of warfarin therapy after the index GIB was estimated for each patient using logistic regression (see the eAppendix for factors included in the propensity score). Cox proportional hazards modeling was used to adjust for potentially confounding factors in the assessment of the association of warfarin resumption with time to thrombosis, recurrent GIB, or death (see the eAppendix for factors included in each model).

To limit the effect of the severity of the index GIB on death, post hoc adjusted hazards modeling on time to death was performed, in which patients who died within 1 week of the index GIB were excluded. In addition, discrete time-varying and categorical variables were constructed based on length of time patients were off warfarin therapy after the index GIB to assess if there was a time-dependent effect of warfarin therapy interruption on the outcomes. Length of warfarin therapy interruption was categorized as 0 days, 1 to 7 days, 8 to 14 days, 15 to 90 days, and warfarin therapy not resumed. Individual post hoc adjusted hazards models for time to recurrent GIB, thrombosis, and death were constructed with the time-varying warfarin exposure variable. Because warfarin therapy was not interrupted in all cases, individual post hoc adjusted hazards models for time to recurrent GIB, thrombosis, and death were constructed with binary and time-varying warfarin exposure variables after removing patients who did not interrupt warfarin therapy and had an index GIB location of rectum-anus (n=24) and then all patients who did not interrupt warfarin therapy (n=41). Post hoc tests of association were performed with the categorical exposure variables and study outcomes. Because adjusted hazard modeling subanalyses using time-varying warfarin exposure and/or removing patient groups who did not have warfarin therapy interruption revealed similar results for the thrombosis, recurrent GIB, and death outcomes, only results of the initial analysis are reported.

Further analyses included comparisons of outcomes and patient characteristics between patients who did and did not experience a recurrent GIB and were and were not dead at the end of follow-up, respectively. Statistical analyses were performed using Intercooled STATA version 9.0 software (StataCorp). The α level was set at .05, and all tests were 2-sided.

Table 1. Baseline Characteristics According to Warfarin Therapy Status Following GIB[a]

Characteristic	Overall Cohort (N = 442)	Resumed Warfarin Therapy (n = 260)	Did Not Resume Warfarin Therapy (n = 182)	*P* Value[b]
Age, mean (SD), y[c]	74.2 (12.1)	71.8 (12.0)	77.7 (11.3)	<.001
Male, No.	222 (50.2)	129 (49.6)	93 (51.1)	.76
INR target[c]				
2.0	17 (3.8)	7 (2.7)	10 (5.5)	.13
2.5	363 (82.4)	203 (78.1)	160 (88.4)	.008
≥3.0	61 (13.8)	50 (19.2)	11 (6.1)	<.001
Chronic Disease Score, mean (SD)[d]	8.4 (3.1)	8.2 (3.0)	8.6 (3.2)	.24
INR at GIB, median (IQR)	3.0 (2.3-4.3)	2.9 (2.3-4.2)	3.2 (2.4-4.5)	.19
Primary indication for anticoagulation therapy[2]				
Atrial fibrillation	223 (50.5)	120 (46.2)	103 (56.6)	.03
Venous thromboembolism[d]	108 (24.4)	67 (25.8)	41 (22.5)	.44
Prosthetic heart valve	42 (9.5)	40 (15.4)	2 (1.1)	<.001
Other	69 (15.6)	33 (12.7)	36 (19.8)	.04
Risk factors[e]				
Alcoholism	4 (0.9)	3 (1.2)	1 (0.6)	.65
Diabetes mellitus	12 (2.7)	6 (2.3)	6 (3.3)	.53
Hypertension	237 (53.6)	128 (49.2)	109 (59.9)	.03
Heart failure	110 (24.9)	63 (24.2)	47 (25.8)	.70
Renal insufficiency	49 (11.1)	26 (10.0)	23 (12.6)	.39
Prior venous thrombosis	70 (15.8)	44 (16.9)	26 (14.3)	.46
Prior arterial thrombosis	1 (0.2)	0	1 (0.6)	.41
Prior ischemic stroke/TIA	39 (8.8)	4 (10.3)	35 (8.6)	.61
Cancer	6 (1.4)	3 (1.2)	3 (1.7)	.69
GIB location				
Large intestine	116 (26.2)	73 (28.1)	43 (23.6)	.26
Mouth-esophagus	30 (6.8)	20 (7.7)	10 (5.5)	.37
Rectum-anus	64 (14.5)	51 (19.6)	13 (7.1)	<.001
Small intestine	14 (3.2)	7 (2.7)	7 (3.9)	.50
Stomach-duodenum	125 (28.3)	65 (25.0)	60 (32.7)	.07
Not identified	93 (21.0)	44 (16.9)	49 (26.9)	.01
Aspirin use, mg				
None	237 (53.6)	145 (55.8)	92 (50.3)	.28
50	2 (0.5)	1 (0.4)	1 (0.6)	.83
81	187 (42.3)	107 (41.2)	80 (44.0)	.56
162	3 (0.7)	1 (0.4)	2 (1.1)	.57
325	13 (2.9)	6 (2.3)	7 (3.9)	.35
Days from warfarin therapy initiation, median (IQR)[c]	891 (167-2477)	1026 (267-2877)	688 (115-2086)	.006
Percentage of INRs in range, mean (SD)[f]	47.7 (27.6)	50.1 (28.4)	44.4 (26.2)	.08
$CHADS_2$ score, median (IQR) [No. of patients][g]	2 (1-2) [223]	2 (1-2) [120]	2 (1-2) [103]	.16

Abbreviations: $CHADS_2$, congestive heart failure, hypertension, age 75 years or older, diabetes mellitus, and prior stroke or TIA; GIB, gastrointestinal tract bleeding; INR, international normalized ratio; IQR, interquartile range; TIA, transient ischemic attack.

[a] Data are given as number (percentage) of patients unless otherwise specified.

[b] Comparison between "resumed warfarin therapy" and "did not resume warfarin therapy" groups.

[c] As of date of the initial GIB event.

[d] Deep vein thrombosis of the upper or lower extremity and pulmonary embolism.

[e] During the 180 days prior to the initial GIB event.

[f] During the 90 days prior to the initial GIB event.

[g] Among patients with atrial fibrillation only.

RESULTS

Of 502 patients identified as having GIB using administrative data, the index GIB was not confirmed during medical record review in 57 patients, and 3 patients were not receiving warfarin. Therefore, 442 patients with warfarin-associated index GIB were included in the analyses (**Table 1**). The mean age was 74.2 years, 50.2% were male, and 46.4% used aspirin at some point during the 90 days prior to the index GIB. Indications for warfarin therapy included the following: prevention of atrial fibrillation–related stroke or systemic embolization (50.5%); treatment or secondary prevention of venous thrombosis (24.4%); and prevention of prosthetic heart valve thromboembolic complications (9.5%). The median (IQR) INR on presentation was 3.0 (2.3-4.3). Approximately one-third of patients (30.5%) were initially treated in the ICU; 24.2% were evaluated and discharged directly from the ED (**Table 2**). Following the index GIB, 260 patients (58.8%) resumed warfarin therapy, including 41 patients whose warfarin therapy was never stopped. Median (IQR) time to resumption of warfarin was 4 days (2-9 days). Prosthetic heart valve indication for warfarin therapy (15.4% vs 1.1%; $P<.001$) and GIB localized to the rectum-anus (representing predominately hemorrhoidal bleeds) (19.6% vs 7.1%; $P<.001$) were more

Table 2. Overall Outcomes Among Warfarin-Treated Patients With GIB[a]

Outcome	Overall Cohort (n = 442)	Resumed Warfarin Therapy (n = 260)	Did Not Resume Warfarin Therapy (n = 182)	*P* Value[b]
Index GIB management				
Warfarin therapy stopped	400 (90.7)	219 (84.2)	182 (100)	<.001
Phytonadione administered	282 (63.8)	157 (60.4)	125 (68.7)	.07
Fresh-frozen plasma provided	211 (47.7)	106 (40.8)	105 (57.7)	<.001
Blood transfusion	252 (57.0)	119 (45.8)	133 (73.1)	<.001
Treated in ED only	107 (24.2)	83 (31.9)	24 (13.2)	<.001
Treated in ICU	135 (30.5)	55 (21.2)	80 (44.0)	<.001
Length of stay, median (IQR), d	3 (1-4)	2 (1-4)	3 (2-5)	<.001
Days to warfarin therapy resumption, median (IQR)[c,d]	4 (2-9)	4 (2-9)	NA	NA
Low-molecular-weight heparin use	44 (10.0)	39 (15.0)	5 (2.8)	<.001
Primary outcomes[e]				
Thrombosis	11 (2.5)	1 (0.4)	10 (5.5)	<.001
Recurrent GIB	36 (8.4)	26 (10.0)	10 (5.5)	.09
Deceased	52 (11.8)	15 (5.8)	37 (20.3)	<.001

Abbreviations: ED, emergency department; GIB, gastrointestinal tract bleeding; ICU, intensive care unit; INR, international normalized ratio; IQR, interquartile range; NA, not applicable.
[a] Data are given as number (percentage) of patients unless otherwise specified.
[b] Comparison between "resumed warfarin therapy" and "did not resume warfarin therapy" groups.
[c] In the 90 days following the initial GIB event but before recurrent GIB, where applicable.
[d] Including only patients who restarted warfarin therapy.
[e] In the 90 days following the initial GIB event.

Table 3. Description of Thrombotic Events

Patient No./ Sex/Age, y	Indication	Resumed Warfarin Therapy	Days From the Index GIB to Thrombosis	Thrombosis Type	Fatal
1/M/74	Deep vein thrombosis	No	90	Pulmonary embolism	No
2/F/85	Atrial fibrillation	No	8	Systemic embolism	No
3/M/75	Deep vein thrombosis	Yes	74	Deep vein thrombosis	No
4/M/85	Atrial fibrillation	No	27	Stroke	No
5/F/84	Atrial fibrillation	No	8	Stroke	Yes
6/M/76	Atrial fibrillation	No	23	Stroke	Yes
7/M/65	Pulmonary embolism	No	39	Pulmonary embolism and deep vein thrombosis	No
8/M/71	Stroke	No	8	Stroke	No
9/F/91	Atrial fibrillation	No	73	Stroke	Yes
10/M/62	Pulmonary embolism	No	13	Pulmonary embolism	No
11/F/95	Pulmonary embolism	No	17	Deep vein thrombosis	No

common among the 260 patients who resumed warfarin therapy compared with those who did not, respectively. In contrast, compared with those resuming warfarin therapy, older patients (mean age, 71.8 years [resumed warfarin therapy] vs 77.7 years [did not resume warfarin therapy]; $P < .001$) and patients for whom the GIB source was not identified (16.9% [resumed warfarin therapy] vs 26.9% [did not resume warfarin therapy]; $P = .01$) were less likely to resume warfarin therapy (Table 1).

90-DAY OUTCOMES: THROMBOSIS

During the 90-day follow-up period, 11 patients (2.5%) experienced a thrombotic event (6 arterial [5 strokes and 1 systemic embolus] and 5 venous [3 pulmonary emboli and 2 deep vein thromboses]), and 3 of the strokes were fatal (**Table 3**). Of the 260 patients who resumed warfarin therapy following the index GIB, 1 (0.4%) had a thrombotic event (deep vein thrombosis) compared with 10 of 182 patients (5.5%) who did not resume warfarin therapy ($P < .001$). Warfarin therapy resumption after the index GIB was associated with a lower risk for thrombosis (hazard ratio [HR], 0.05; 95% CI, 0.01-0.58) in a multivariable analysis that controlled for the propensity score, CDS, age, and sex (**Figure**, A). For patients resuming warfarin therapy, thrombosis rates were similar regardless of the duration of warfarin therapy interruption. Patients who either never interrupted warfarin therapy or resumed therapy within 14 days of the index GIB experienced no thromboses.

90-DAY OUTCOMES: RECURRENT GIB

Of the 442 patients, 36 (8.4%) had recurrent GIB (Table 2). Compared with those who did not resume warfarin therapy, a numerically higher proportion of patients resuming warfarin therapy had recurrent GIB, but

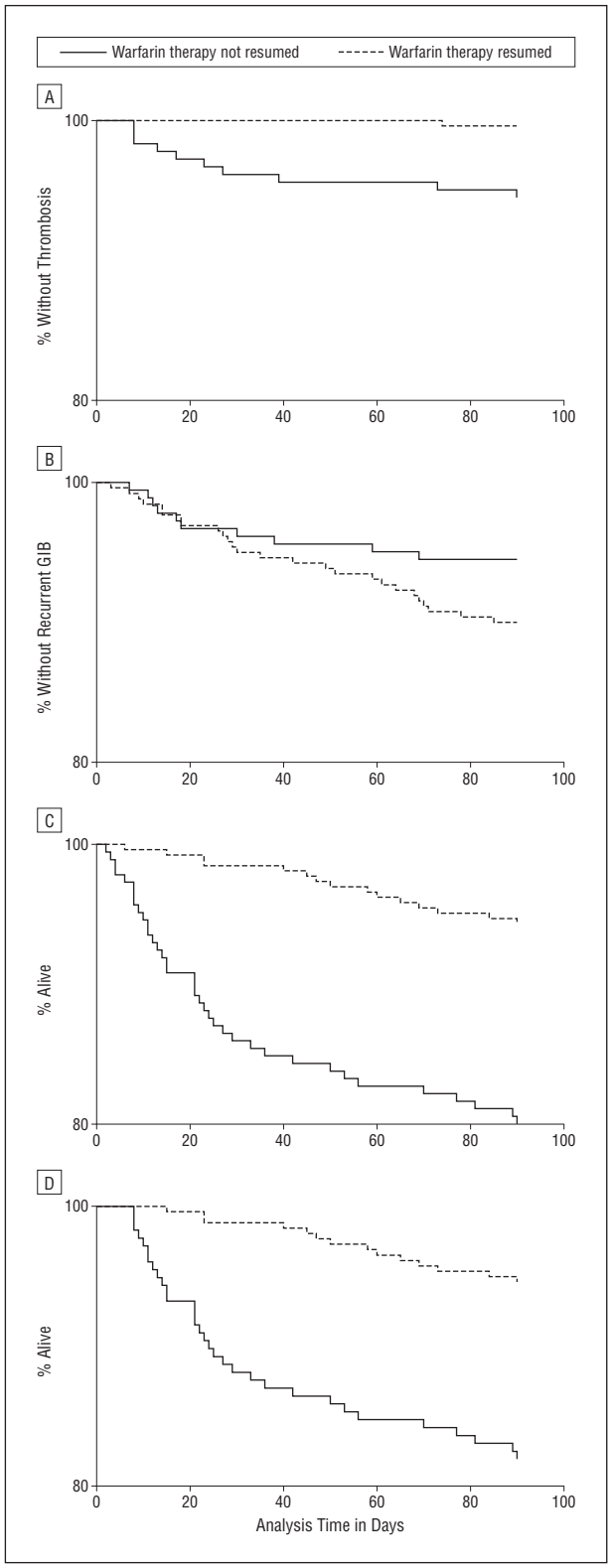


Figure. Time-to-outcome analysis according to resuming warfarin therapy status. A, Thrombosis ($P = .002$, log-rank test); B, recurrent gastrointestinal tract bleeding (GIB) ($P = .10$, log-rank test); C, death ($P < .001$, log-rank test); and D, death including only patients who died at least 7 days after the index GIB ($P < .001$, log-rank test).

this difference was not statistically significant (10.0% [resumed warfarin therapy] vs 5.5% [did not resume warfarin therapy]; $P = .09$). Multivariable analysis that controlled for the propensity score, CDS, age, sex, indication for warfarin use, prior heart failure diagnosis, location of GIB, pre-GIB target INR, pre-GIB percentage of INRs in range, reception of low-molecular-weight heparin, length of ED/inpatient stay, and acute GIB treatment (blood transfusion) also revealed that the risk for recurrent GIB associated with warfarin therapy resumption was not increased significantly (HR, 1.32; 95% CI, 0.50-3.57) (Figure, B). Compared with all other patients, the rate of recurrent GIB was significantly increased when warfarin therapy was resumed between 1 and 7 days after the index GIB (6.23% vs 12.4%, respectively; $P = .03$). Although 5 of the index GIB events were eventually fatal, no recurrent GIB resulted in death. The median (IQR) time from warfarin therapy resumption to recurrent GIB was 27 days (11-58 days). There was no association between more aggressive management of the index GIB (eg, ICU admission, use of blood products) and recurrent GIB (all $P > .05$) (**Table 4**).

90-DAY OUTCOMES: DEATH

During the 90-day follow-up period, 52 patients (11.8%) died (**Table 5**). The most common causes of death were related to malignancy (28.8%), infection (19.2%), and cardiac disease (17.3%). No deaths were attributed to recurrent GIB. Compared with survivors, patients who died were older ($P = .03$) and had higher CDS ($P = .004$). Patients with an index GIB localized to the mouth-esophagus died less frequently, and those with an index GIB with an unidentified bleeding source died more frequently. Warfarin therapy resumption after the index GIB was associated with a lower risk for death (HR, 0.31; 95% CI, 0.15-0.62) in multivariable analysis that controlled for the propensity score, CDS, age, sex, location of GIB, ICU admission, hypertension, prior stroke diagnosis, pre-GIB percent of INRs in range, reception of low-molecular-weight heparin, length of ED/inpatient stay, and acute GIB treatment (blood transfusion) (Figure, C). This strong association persisted in a post hoc analysis excluding all patients who died within 1 week of the index GIB (Figure, D). The death rate during follow-up was lowest when warfarin therapy was resumed between 15 and 90 days after the index GIB (2.3%, $P = .04$ compared with all other patients).

COMMENT

Gastrointestinal tract bleeding is a common complication of warfarin therapy. This retrospective cohort study evaluated 90-day outcomes among warfarin-treated patients with GIB. The results highlight the clinical dilemma of managing warfarin therapy following a hospitalization or ED visit for GIB. Although we observed a numerical increase in recurrent GIB associated with not interrupting or resuming warfarin therapy in the 90 days after the index GIB, this increase was not statistically significant. However, a decision not to resume warfarin therapy was associated with a significantly increased risk for both thrombosis and death from any cause. Furthermore, while no GIB recurrences were fatal, 3 patients with atrial fibrillation had fatal strokes during warfarin therapy

Table 4. Outcomes According to Recurrent GIB Status[a]

Outcome	Overall Cohort (N = 442)	Recurrent GIB (n = 36)	No Recurrent GIB (n = 406)	*P* Value[b]
Index GIB management				
Warfarin therapy stopped[c]	400 (90.7)	33 (91.7)	367 (90.4)	.79
Phytonadione administered[c]	282 (63.8)	26 (72.2)	256 (63.1)	.27
Fresh-frozen plasma provided[c]	211 (47.7)	19 (52.8)	192 (47.3)	.53
Blood transfusion[c]	252 (57.0)	25 (69.4)	227 (55.9)	.12
Treated in ED[c]	107 (24.2)	10 (27.8)	97 (23.9)	.60
Treated in ICU[c]	135 (30.5)	14 (38.9)	121 (29.8)	.26
Warfarin therapy resumed[d]	260 (58.8)	26 (72.2)	234 (57.6)	.09
Length of stay, median (IQR), d	3 (1-4)	3 (2-4)	2 (1-4)	.45
Days to warfarin therapy resumption, median (IQR)[d,e]	4 (2-9)	3 (1-6)	4 (2-9)	.41
Low-molecular-weight heparin use[f]	44 (10.0)	6 (16.7)	38 (9.4)	.16
Primary outcomes				
Thrombosis[f]	11 (2.5)	0	11 (2.7)	.39
Deceased[f]	52 (11.8)	2 (5.6)	50 (12.3)	.23

Abbreviations: ED, emergency department; GIB, gastrointestinal tract bleeding; ICU, intensive care unit; INR, international normalized ratio; IQR, interquartile range.

[a] Data are given as number (percentage) of patients unless otherwise specified.

[b] Comparison between "recurrent GIB" and "no recurrent GIB" groups.

[c] At time of the initial GIB event.

[d] In the 90 days following the initial GIB event but before recurrent GIB, where applicable.

[e] Including only patients who restarted warfarin therapy.

[f] In the 90 days following the initial GIB event.

discontinuation following a GIB event. While the increased risk of thrombosis and death associated with any warfarin therapy interruption has been reported previously in a Danish registry of patients with atrial fibrillation,[14] to our knowledge, ours is the first study to report this observation in a cohort of patients receiving warfarin for diverse indications specifically in the context of recent GIB. In our study, the exact date and duration of warfarin therapy interruption and adverse events were verified through medical record review, whereas in the Danish study,[14] the date of warfarin therapy interruption was estimated from warfarin prescription claims data, the reasons why patients interrupted therapy were unknown, and adverse events were not confirmed by medical record review.

The theoretical concern that abrupt warfarin therapy discontinuation following GIB causes a temporary hypercoagulable state may be relevant to the observed increase in thrombosis in patients who did not resume warfarin therapy,[15] although no thrombotic events occurred within 7 days of warfarin therapy interruption. It is difficult, if not impossible, to determine the time course between thrombus formation and subsequent clinical manifestations. However, the laboratory evidence supporting the actual existence of "rebound hypercoagulability" is inconsistent, and clinical trials have failed to demonstrate increased thromboembolic risk associated with abrupt anticoagulant withdrawal.[15]

Our observation that a decision not to resume warfarin therapy is associated with higher overall mortality was unexpected and not readily explained, given that only 3 of the 37 deaths in the group not resuming warfarin therapy were attributed to thrombosis. We attempted to control for possible confounding of the warfarin therapy resumption indicator by including pertinent factors in multivariable analysis and by performing propensity score analysis; however, the association persisted. It is possible that patients with a more serious index GIB (who would presumably be more likely to die) were also less likely to resume anticoagulation. However, the association between resuming warfarin therapy and lower mortality persisted with modeling that adjusted for ICU admission as well as blood transfusions—interventions that would be expected to be markers of a more serious initial GIB. To further explore explanations for the association between a decision not to resume warfarin therapy and death, we reanalyzed the data after excluding patients who died during the first week after the index GIB because these patients would have had less opportunity to resume warfarin therapy. Despite this, we found that the association remained significant. We acknowledge that residual confounding was likely present despite rigorous efforts at mitigation through various analytical approaches. Therefore, the apparent increase in nonthrombotic deaths when warfarin therapy was not resumed may suggest that the treating physicians were reluctant to resume warfarin therapy in sicker patients with a higher risk of death in general.

Our results provide some guidance regarding the optimal timing of warfarin therapy resumption following GIB, but clinical judgment remains a critical factor in this difficult decision. Resumption of warfarin therapy between days 1 and 7 following a GIB event was associated with a higher risk of recurrent GIB but lower risk of thrombosis. A better understanding of the propensity for recurrent hemorrhage and its severity across the spectrum of anatomic lesions would help to inform the decision of optimal timing of anticoagulation resumption, an issue of major importance for individuals at highest risk of thromboembolism.

Our study is limited in that we used data from administrative databases, and thus not all factors that affect clini-

Table 5. Patient Characteristics by Death Status at End of Follow-up[a]

Characteristic	Overall Cohort (n = 442)	Died (n = 52)	Alive (n = 390)	*P* Value[b]
Age, mean (SD), y[c]	74.2 (12.1)	77.6 (11.0)	73.8 (12.2)	.03
Male	222 (50.2)	29 (55.7)	193 (49.5)	.40
INR target[c]				
2.0	17 (3.8)	2 (3.9)	15 (3.8)	.61
2.5	363 (82.4)	42 (80.8)	321 (82.5)	.79
≥3.0	61 (13.8)	8 (15.3)	53 (13.7)	.73
Chronic Disease Score, mean (SD)[d]	8.4 (3.1)	9.5 (3.3)	8.3 (3.0)	.004
INR at GIB, median (IQR)	3.0 (2.3-4.3)	3.3 (2.1-5.7)	3.0 (2.3-4.2)	.42
Primary indication for anticoagulation therapy[2]				
Atrial fibrillation	223 (50.5)	27 (51.9)	196 (50.3)	.82
Venous thromboembolism	108 (24.4)	10 (25.7)	98 (25.1)	.35
Prosthetic heart valve	42 (9.5)	4 (7.7)	38 (9.7)	.64
Other	69 (15.6)	11 (21.2)	58 (14.9)	.24
Risk factors[d]				
Alcoholism	4 (0.9)	0	4 (1.0)	.61
Diabetes mellitus	12 (2.7)	2 (3.9)	10 (2.6)	.59
Hypertension	237 (53.6)	33 (63.5)	204 (52.3)	.13
Heart failure	110 (24.9)	11 (21.2)	99 (25.4)	.51
Renal insufficiency	49 (11.1)	4 (7.7)	45 (11.5)	.49
Prior venous thrombosis	70 (15.8)	8 (15.4)	62 (15.9)	.92
Prior arterial thrombosis	1 (0.2)	1 (1.9)	0	.12
Prior ischemic stroke/TIA	39 (8.8)	9 (17.3)	30 (7.7)	.02
Cancer	6 (1.4)	1 (1.9)	5 (1.3)	.53
GIB location				
Large intestine	116 (26.2)	18 (34.6)	98 (25.1)	.14
Mouth-esophagus	30 (6.8)	0	30 (7.7)	.04
Rectum-anus	64 (14.5)	5 (9.6)	59 (15.1)	.29
Small intestine	14 (3.2)	1 (1.9)	13 (3.3)	.83
Stomach-duodenum	125 (28.3)	10 (19.2)	115 (29.5)	.12
Not identified	93 (21.0)	18 (34.6)	75 (19.2)	.01
Aspirin dose, mg[e]				
None	237 (53.6)	26 (50.0)	211 (54.1)	.58
50	2 (0.5)	0	2 (0.5)	.78
81	187 (42.3)	22 (42.3)	165 (42.3)	>.99
162	3 (0.7)	2 (3.9)	1 (0.3)	.04
325	13 (2.9)	2 (3.4)	11 (2.8)	.66
Days from warfarin therapy initiation, median (IQR)[c]	891 (167-2477)	988 (162-2450)	886 (167-2477)	.72
$CHADS_2$ score, median (IQR) [No. of patients][f]	2 (1-2) [223]	2 (1-2) [27]	2 (1-2) [197]	.21

Abbreviations: $CHADS_2$, congestive heart failure, hypertension, age 75 years or older, diabetes mellitus, and prior stroke or TIA; GIB, gastrointestinal tract bleeding; INR, international normalized ratio; IQR, interquartile range; TIA, transient ischemic attack.

[a] Data are given as number (percentage) of patients unless otherwise specified.

[b] Comparison between "dead" and "alive" groups.

[c] As of the date of the initial GIB event.

[d] During the 180 days prior to the initial GIB event.

[e] During the 90 days prior to the initial GIB event.

[f] Among patients with atrial fibrillation only.

cal decision making could be collected. However, the confirmation of thrombosis, recurrent GIB, and death outcomes via medical chart and death certificate review strengthens the validity of our results. The observed results are biologically plausible because the thrombotic events correlated with the indication for warfarin therapy—patients with atrial fibrillation experienced strokes or systemic embolus, while patients with venous thromboembolism had recurrent venous thromboembolism (Table 2). Despite potential confounding, our study likely underestimates the strength of the association between thrombosis and not resuming warfarin therapy because patients at the highest risk of thrombosis (eg, presence of mechanical heart valve, high $CHADS_2$ score) were probably less likely to remain off warfarin therapy following GIB. Similarly, the strength of association between recurrent GIB and resuming warfarin therapy may be underestimated because patients perceived at high risk for further GIB probably were less likely to resume warfarin therapy. Accurate recording of baseline aspirin use status was facilitated by the records maintained by CPAS. However, aspirin use status was not routinely documented in index GIB discharge summaries, and a sizeable proportion of patients did not resume warfarin therapy and were thus not followed by CPAS after the index GIB. Therefore, we were not able to accurately record aspirin use following the index GIB and acknowledge that lack of information on post-GIB aspirin use and its potential influence on the risk of recurrent GIB, thrombosis, and death is a limitation.

Our study shows that the decision to not resume warfarin therapy in the 90 days following GIB is associated

with increased risk for thrombosis and death. Our analysis suggests that, for many patients who have experienced GIB, the benefits of resuming warfarin therapy will outweigh the risks. Further research will be needed to identify the optimal duration of warfarin interruption after a GIB event and the patients for whom a more prolonged interruption can be justified.

Accepted for Publication: June 4, 2012.
Published Online: September 17, 2012. doi:10.1001/archinternmed.2012.4261
Author Affiliations: Clinical Pharmacy Research Team (Drs Witt and Delate) and Clinical Pharmacy Anticoagulation & Anemia Service (Dr Clark), Kaiser Permanente Colorado, Aurora; University of Colorado Skaggs School of Pharmacy and Pharmaceutical Sciences, Denver (Drs Witt, Delate, and Clark); Department of Internal Medicine, University of New Mexico School of Medicine, Albuquerque (Dr Garcia); Boston University School of Medicine, Boston, Massachusetts (Dr Hylek); Department of Clinical and Experimental Medicine, University of Insubria, Varese, Italy (Drs Ageno and Dentali); and Department of Medicine, McMaster University, Hamilton, Ontario, Canada (Dr Crowther).
Correspondence: Daniel M. Witt, PharmD, FCCP, BCPS, Clinical Pharmacy Anticoagulation Service, Kaiser Permanente of Colorado, 16601 E Centretech Pkwy, Aurora, CO 80011 (dan.m.witt@kp.org).
Author Contributions: The principal investigator, Dr Witt, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Witt, Delate, Garcia, Clark, and Dentali. *Acquisition of data:* Witt, Delate, Garcia, and Clark. *Analysis and interpretation of data:* Witt, Delate, Garcia, Hylek, Ageno, and Dentali. *Drafting of the manuscript:* Witt, Delate, Garcia, and Dentali. *Critical revision of the manuscript for important intellectual content:* Witt, Garcia, Clark, Hylek, Ageno, and Dentali. *Statistical analysis:* Witt and Delate. *Obtained funding:* Witt and Garcia. *Administrative, technical, and material support:* Garcia and Clark. *Study supervision:* Witt and Ageno.
Financial Disclosure: Dr Garcia has served as an advisor to Boehringrer Ingelheim, Bristol-Meyers Squibb, Daiichi Sankyo, and CSL Behring. Dr Hylek has served as an advisor to Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Daiichi Sankyo, Johnson and Johnson, and Pfizer and participated in clinical symposia sponsored by Bayer, Boehringer Ingelheim, and Bristol-Myers Squibb. Dr Crowther has served as an advisor to Leo Pharma, Pfizer, Bayer, Boehringer Ingelheim, Alexion, CSL Behring, and Artisan Pharma; has prepared educational materials for Pfizer, Octapharm, and CSL Behring; has provided expert testimony for Bayer; and holds a Career Investigator award from the Heart and Stroke Foundation of Ontario and the Leo Pharma Chair in Thromboembolism Research at McMaster University. Dr Crowther's affiliated institution (McMaster University) has received funding for research projects from Boehringer Ingelheim, Octapharm, Pfizer, and Leo Pharma.
Funding/Support: CSL Behring LLC provided funding for the study.
Role of the Sponsor: CSL Behring LLC was not involved in the study design, collection, analysis, or interpretation of the data or in preparation and submission of the manuscript for publication.
Online-Only Material: The eAppendix is available at http://www.archinternmed.com.
Additional Contributions: Aubrey E. Jones assisted with the extraction of information from medical records.

REFERENCES

1. Schelleman H, Brensinger CM, Bilker WB, Hennessy S. Antidepressant-warfarin interaction and associated gastrointestinal bleeding risk in a case-control study. *PLoS One*. 2011;6(6):e21447.
2. White RH, McKittrick T, Takakuwa J, Callahan C, McDonell M, Fihn S; National Consortium of Anticoagulation Clinics. Management and prognosis of life-threatening bleeding during warfarin therapy. *Arch Intern Med*. 1996;156(11):1197-1201.
3. Guerrouij M, Uppal CS, Alklabi A, Douketis JD. The clinical impact of bleeding during oral anticoagulant therapy: assessment of morbidity, mortality and post-bleed anticoagulant management. *J Thromb Thrombolysis*. 2011;31(4):419-423.
4. Kearon C, Hirsh J. Management of anticoagulation before and after elective surgery. *N Engl J Med*. 1997;336(21):1506-1511.
5. Kuwada SK, Balm R, Gostout CJ. The risk of withdrawing chronic anticoagulation because of acute GI bleeding. *Am J Gastroenterol*. 1996;91(6):1116-1119.
6. Nieto JA, Bruscas MJ, Ruiz-Ribo D, et al; RIETE Investigators. Acute venous thromboembolism in patients with recent major bleeding: the influence of the site of bleeding and the time elapsed on outcome. *J Thromb Haemost*. 2006;4(11):2367-2372.
7. Claassen DO, Kazemi N, Zubkov AY, Wijdicks EF, Rabinstein AA. Restarting anticoagulation therapy after warfarin-associated intracerebral hemorrhage. *Arch Neurol*. 2008;65(10):1313-1318.
8. Phan TG, Koh M, Wijdicks EF. Safety of discontinuation of anticoagulation in patients with intracranial hemorrhage at high thromboembolic risk. *Arch Neurol*. 2000;57(12):1710-1713.
9. Witt DM, Sadler MA, Shanahan RL, Mazzoli G, Tillman DJ. Effect of a centralized clinical pharmacy anticoagulation service on the outcomes of anticoagulation therapy. *Chest*. 2005;127(5):1515-1522.
10. Clark DO, Von Korff M, Saunders K, Baluch WM, Simon GE. A chronic disease score with empirically derived weights. *Med Care*. 1995;33(8):783-795.
11. Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol*. 1992;45(2):197-203.
12. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA*. 2001;285(22):2864-2870.
13. D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998;17(19):2265-2281.
14. Raunso J, Selmer C, Olesen JB, et al. Increased short-term risk of thromboembolism or death after interruption of warfarin treatment in patients with atrial fibrillation [published online December 23, 2011]. *Eur Heart J*. doi:10.1093/eurheartj/ehr454.
15. Holbrook A, Schulman S, Witt DM, et al; American College of Chest Physicians. Evidence-based management of anticoagulant therapy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141(2)(suppl):e152S-e184S.

BMJ open accessible medical research

PRESS RELEASE

Gender as risk factor for 30 days post-discharge hospital utilisation: a secondary data analysis

Shaula Woz,[1] Suzanne Mitchell,[1] Caroline Hesko,[1] Michael Paasche-Orlow,[2] Jeffrey Greenwald,[2] V K Chetty,[1] Julie O'Donnell,[1] Brian Jack[1]

To cite: Woz S, Mitchell S, Hesko C, *et al*. Gender as Risk factor for 30 days post-discharge hospital utilisation: a secondary data analysis. *BMJ Open* 2012;**2**: e000428. doi:10.1136/bmjopen-2011-000428

► Prepublication history and additional appendices for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2011-000428).

SM is a consultant and speaker for the non-product health topics series on relationship centred care for Merck & Co.

Received 4 October 2011
Accepted 12 March 2012

[1]Department of Family Medicine, Boston University School of Medicine, Boston, Massachusetts, USA
[2]Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

Correspondence to
Dr Suzanne Mitchell;
suzanne.mitchell@bmc.org

ABSTRACT

Objective: In the 30 days after hospital discharge, hospital utilisation is common and costly. This study evaluated the association between gender and hospital utilisation within 30 days of discharge.

Design: Secondary data analysis using Poisson regression stratified by gender.

Participants: 737 English-speaking hospitalised adults from general medical service in urban, academic safety-net medical centre who participated in the Project Re-Engineered clinical trial (clinicaltrials.gov identifier: NCT00252057).

Main outcome measure: The primary end point was hospital utilisation, defined as total emergency department visits and hospital readmissions within 30 days after index discharge.

Results: Female subjects had a rate of 29 events for every 100 people and male subjects had a rate of 47 events for every 100 people (incident rate ratio (IRR) 1.62, 95% CI 1.28 to 2.06). Among men, risk factors included hospital utilisation in the 6 months prior to the index hospitalisation (IRR 3.55, 95% CI 2.38 to 5.29), being unmarried (IRR 1.72, 95% CI 1.12 to 2.64), having a positive depression screen (IRR 1.53, 95% CI 1.09 to 2.13) and no primary care physician (PCP) visit within 30 days (IRR 1.64, 95% CI 1.08 to 2.50). Among women, the only risk factor was hospital utilisation in the 6 months prior to the index hospitalisation (IRR 3.08, 95% CI 1.86 to 5.10).

Conclusions: In our data, male subjects had a higher rate of hospital utilisation within 30 days of discharge than female subjects. For men—but not for women—risk factors were being retired, unmarried, having depressive symptoms and having no PCP visit within 30 days. Interventions addressing these factors might lower hospital utilisation rates observed among men.

ARTICLE SUMMARY

Article focus

- Early hospital readmission is a common and costly occurrence in the USA. Men often use hospital emergency departments for usual source of medical care.
- We aimed to study whether men are therefore more likely to be readmitted to the hospital within 30 days of an index discharge.

Key messages

- Men have higher rates of 30-day readmission to hospital than women in this study group. Men also were less likely to complete a follow-up appointment with their primary care physician after discharge.
- Interventions that promote connecting men to primary care, address social isolation and screen for depressive symptoms may reduce the risk for early readmission among men.

Strengths and limitations of this study

- This analysis was conducted with the Project RED data set, which included, and allowed controlling for certain clinical and social confounders in our analysis such as subjects' comorbidity burden, depression symptoms, homelessness, substance abuse and other similar risk factors, in our analysis.
- This study is limited in that it was conducted at an urban safety-net hospital and may not be generalisable to other types of hospital systems.
- We also relied on subject self-report of any rehospitalisation events outside of the study site, however, were able to confirm 91% of all events using our hospital electronic medical record system.

INTRODUCTION

Hospital utilisation in the 30 days after discharge is costly and may be a marker of poor quality of care. In 2004, the cost for hospital readmissions among Medicare recipients was estimated to be US $17.4 billion.[1] Accordingly, the Affordable Care Act includes multiple provisions designed to improve care transitions. The act includes both funding to stimulate hospitals and community-based providers to coordinate post-discharge services and a programme to withhold payments, of progressively increasing amounts, to hospitals that demonstrate higher rates of readmission within 30 days after discharge.[2] The extent to

which readmissions are preventable is debated.[3 4] However, given the magnitude of the problem, even a moderate reduction in unnecessary readmissions could have a large economic impact. In fact, we have shown that hospital readmissions can be moderately reduced.[5]

With prospective identification of patients at high risk of being readmitted, providers could potentially direct resources to prevent the readmission. In order to fulfil this goal in an efficient and effective manner, providers will need to be able to identify the patients with high risk and may need to understand how to adapt services according to the risk factors identified.

Several factor associated with 30-day rehospitalisation have been well characterised. These include older age,[6 7] comorbidity,[7] income level,[8] history of prior hospitalisation,[6 7] increased length of stay in the index hospitalisation,[9] minority ethnicity,[7 10] depressive symptoms,[11 12] alcohol and drug use[10] and specific clinical conditions (eg, congestive heart failure).[6 7] The role of male gender as a risk factor for post-discharge hospital utilisation has been noted in several disease-specific contexts; yet, to our knowledge, no study of post-discharge hospital utilisation has focused on gender. While gender is not typically considered a modifiable factor, established patterns of health service utilisation that are associated with gender (ie, lower rates of preventive care and fewer visits to primary care among men) may put men at higher risk for poor outcomes after hospital discharge.[13]

Therefore, we conducted a secondary analysis of the Re-engineered Discharge (RED) clinical trial data set to assess the association between gender and the rate of post-discharge hospital utilisation among a cohort of adult patients hospitalised in an urban safety-net hospital. In addition, we sought to identify potential factors contributing to gender-based differences.

METHODS

A full description of the methods for the Project RED trial has been described previously.[5] Briefly, the Project RED trial was a two-armed randomised controlled trial of English-speaking adult patients, 18 years or older, admitted to the teaching service of Boston Medical Center. Seven hundred and forty-nine subjects were enrolled and randomised: 376 in the usual care arm and 373 in the intervention arm. Patients had to have a telephone, be able to comprehend study details and the consent process in English and have plans to be discharged to a US community. Patients were not enrolled if they were admitted from a skilled nursing facility or other hospital, transferred to a different hospital service, admitted for a planned hospitalisation, on hospital precautions, on suicide watch, deaf or blind. A total of 3873 were assessed for eligibility. Due to a lack of available research staff, 1616 patients were not assessed. Of those assessed for eligibility, 1049 did not meet eligibility criteria, 120 were previously enrolled, 527 refused to participate, 474 were unavailable in their hospital room at the time of enrolment and 954 were not approached because the maximum enrolled subject number was reached that day. Seven hundred and forty-nine subjects were enrolled and randomised: 376 in the usual care arm and 373 in the intervention arm. The Institutional Review Board of Boston University approved all study activities. Baseline demographic and clinical characteristics were similar across the study arms.

Primary outcome

The primary outcome was the rate of post-discharge hospital utilisation, defined as the total number of emergency department (ED) visits and readmissions per subject, within 30 days of their index discharge. Any ED visit in which a subject was subsequently admitted to the hospital was only counted as a readmission.

Primary independent variable and covariates

The primary independent variable, gender (male or female), was defined by the hospital electronic medical record (EMR). Potential confounders were identified a priori from the literature on factors associated with post-discharge hospital utilisation and gender including age, marital status, health literacy score,[14] Charlson score,[15] insurance type, employment status, income level, homeless status, hospital utilisation within the 6 months prior to the index hospitalisation, educational attainment, length of hospital stay, race/ethnicity, depressive symptoms and Project RED study group assignment.

Data collection

Outcome data were collected by Project RED research staff, blinded to group assignment, by review of the hospital's EMRs and by contacting subjects by telephone 30 days post-discharge. Dates of subsequent ED visits and readmissions at Boston Medical Center were obtained from the EMRs, while those at other hospitals were collected through subject report. Those subjects who could not be reached within 60 days post-discharge were assumed alive and hospital EMRs were relied upon for primary outcomes. Randomisation provided a balanced study sample, with an equal proportion of male and female participants assigned to each arm of the original trial. Of 749 subjects who had participated in a randomised clinical trial (Project RED), 737 participants were included in this secondary analysis and 12 subjects were removed due to death prior to index discharge,[6] requested removal,[13] previously enrolled[1] and missing data.[1] Data for selected covariates were collected by self-report (age, race, income, marital status, education attainment, employment status, insurance type, homelessness) or using validated tools (health literacy, depressive symptoms) or EMR (length of stay, prior utilisation, Charlson score).

Statistical analysis

Socio-demographic and clinical characteristics of the subjects were compared by gender. Bivariate analyses were conducted to identify gender differences and

potential confounders between gender and post-discharge hospital utilisation within 30 days of index discharge. χ^2 Tests were utilised for categorical variables and t tests for continuous variables. A Poisson regression was conducted using relevant potential confounders to construct the final best-fit model determining the strength of association between gender and hospital utilisation. Several interaction terms in our initial Poisson regression were significant. We therefore decided to evaluate potential interactions between gender and hospital utilisation using a stratified Poisson regression analysis.

Age, length of stay and Charlson score were used as continuous variables. Gender (male or female), marital status (married, not married) and homelessness (homeless within the last 3 months) were treated as dichotomous variables. Categorical variables were created for prior hospital utilisation (no prior visits, one to two prior visits or three or more prior visits in the previous 6 months), educational attainment (less than high school graduate, high school graduate or GED or any college), insurance type (Medicare, Medicaid, private insurance or Massachusetts State Subsidised Free Care), income level (no income, <$10 000/year, $10 000–20 000, $20 000 or more or declined to answer), level of health literacy (grade 3 and below, grades 4–6, grades 7–8 or grade 9 and above) according to the Rapid Estimate of Adult Literacy in Medicine and employment status (employed, not employed, disabled, retired or other).

Hospital utilisation is defined as the sum of emergency room visits and hospitalisations (an emergency room visit that leads to a hospitalisation is counted only as a hospitalisation). Hospital utilisation incidence rates were calculated as the number of hospital utilisation events within 30 days of discharge per subject. Person-time was measured in months. The unadjusted incident rate ratio (IRR) was calculated as the ratio of the rate of hospital utilisation among male patients versus female patients using Poisson regression. p Values and CIs were corrected for over dispersion if necessary.

Poisson models were used to test for statistically significant differences in the number of post-discharge hospital utilisation events at 30 days. Two-sided significance tests were used. p Values of <0.05 were considered to indicate statistical significance. A Kaplan–Meier survival curve was generated for the time to multiple hospital utilisation events for the 30-day period following the index discharge and compared using a log-rank test. All data were analysed with S-Plus 8.0.

RESULTS

The analytic cohort included 367 male subjects and 370 female subjects. Socio-demographic, healthcare utilisation and health status indicator variables, stratified by gender, are shown in table 1. Gender differences existed among a number of variables. For example, male subjects were approximately 4 years younger and were more likely to be white non-Hispanic or Hispanic and less likely black non-Hispanic than female subjects. Men reported a relatively higher income, with almost 10% more men reporting an annual personal income of $20 000 or more compared with women. Male subjects were more likely to have private insurance, while female subjects were more likely to have Medicaid. Women reported having a primary care physician (PCP) at baseline at a significantly higher rate than men (88% vs 74%, p<0.001). Women also had higher levels of depressive symptoms (Patient Health Questionnaire-9, 2.5 vs 1.9), were more likely to report a history of having been diagnosed as having depression (46% vs 26%) and were more likely to report currently taking medicine for depression (26% vs 13%).

Hospital utilisation

Female subjects had a rate of 29 events per 100 people per month and male subjects had a rate of 47 events per 100 people per month (IRR 1.62, 95% CI 1.28 to 2.06) (table 2). This difference is largely attributable to a higher rate of ED visits among male subjects (IRR 2.04, 95% CI 1.45 to 2.86). Furthermore, Kaplan–Meier survival curves for the time to multiple hospital utilisation events in the 30 days following index discharge showed that men were more likely to return to the hospital (p=0.04) (figure 1). At the 30-day follow-up telephone call, fewer men reported understanding their appointments after leaving the hospital compared with women (78% and 87%, p=0.005, respectively) (table 2). In addition, at 30 days post-discharge, women reported visiting their PCPs at a higher rate within the 30 days after their hospital discharge (57% and 49%, p=0.04, respectively).

Risk factors for hospital reutilisation by gender

A Poisson regression that is stratified by gender is shown in table 3. The model is controlled for age, previous hospital visits, employment, marital status, depression, study group, having a PCP at baseline and attending a PCP appointment. Among women, the only predictive factor was hospital utilisation in the 6 months prior to the index hospitalisation. Prior hospitalisation was also a risk factor for returning to the hospital within 30 days among men; however, additional significant factors were (1) being retired, (2) not married, (3) having a positive depression screen, (4) reporting no PCP visit within 30 days and (5) not being reached for the follow-up call at 30 days.

DISCUSSION

Among our subjects, we found that men have a higher rate of hospital utilisation within 30 days of hospital discharge than women. ED visits accounted for most of this difference. Among both men and women, prior hospital utilisation is predictive of future utilisation; however, risk factors including being retired, unmarried and having a positive depression screen were identified as risk factors exclusively in men. Additionally, men fared more poorly at understanding and attending their follow-up appointments, which also appeared to be an

Table 1 Baseline characteristics of subjects by gender*

Characteristics	Male (n=367)	Female (n=370)	p Value
Socio-demographics			
Study arm, intervention	195 (53)	174 (47)	0.10
Age, mean (SD), years	47.9 (14.5)	51.6 (15.5)	<0.01
Race, n (%)			
White non-Hispanic	113 (34)	92 (27)	0.02
Black non-Hispanic	171 (51)	215 (62)	
Hispanic	44 (13)	30 (9)	
Other race or mixed race	7 (2)	9 (3)	
Annual personal income, n (%)			
No income	54 (15)	45 (12)	<0.01
<$10 000	53 (15)	83 (22)	
$10 000–$19 999	59 (16)	71 (19)	
$20 000 or more	102 (28)	72 (19)	
Declined to answer	97 (27)	99 (27)	
Health insurance, n (%)			
Private	66 (18)	53 (14)	<0.01
Medicaid	157 (43)	199 (54)	
Medicare	46 (13)	52 (14)	
Free care†	93 (26)	64 (17)	
Highest educational level, n (%)			
Some high school	97 (26)	91 (25)	0.75
High school graduate or GED	136 (37)	146 (40)	
Any college	134 (37)	131 (36)	
REALM health literacy score, mean (SD)	48.1 (21.4)	49.5 (21.1)	0.35
Health literacy level‡, n (%)			
Grade 3 and below	55 (15)	52 (15)	0.31
Grades 4–6	45 (13)	31 (9)	
Grades 7–8	107 (30)	121 (34)	
Grade 9 and above	149 (42)	154 (43)	
Current marital status, n (%)			
Married	112 (31)	116 (31)	0.83
Not married	253 (69)	253 (69)	
Current employment status, n (%)			
Employed	149 (41)	116 (31)	<0.01
Unemployed	72 (20)	70 (19)	
Disabled	78 (21)	86 (23)	
Retired	58 (16)	73 (20)	
Other (student, homemaker, other)	8 (2)	24 (7)	
Homeless in last 3 months, n (%)	40 (11)	35 (10)	0.52
Healthcare utilisation			
Hospital utilisation (hospital utilisations in 6 months prior to index admission)			
0	165 (45)	143 (39)	0.23
1–2	125 (34)	143 (39)	
≥3	77 (21)	82 (22)	
Prior hospital admissions, mean (SD)§	0.66 (1.4)	0.66 (1.1)	0.97
Prior ED visits, mean (SD)§	0.9 (1.8)	1.0 (1.6)	0.40
Length of stay, mean (SD), days	2.7 (3.6)	2.6 (2.5)	0.51
Had PCP at enrolment, n (%)	271 (74)	324 (88)	<0.01
Health status indicators			
Charlson comorbidity score,¶ mean (SD)	1.1 (1.9)	1.4 (1.9)	0.05
SF-12,** mean (SD)			
PCS	40.7 (7.0)	40.2 (7.6)	0.30
MCS	46.6 (9.2)	46.3 (9.9)	0.67
PHQ-9 depression screen,†† mean (SD)	1.9 (2.2)	2.5 (2.3)	<0.01
PHQ-9 depression screen,†† n (%)			
Major depressive disorder, n (%)	48 (13)	72 (19)	0.02
Other depressive disorder, n (%)	56 (15)	60 (16)	0.76
Any depressive disorder, n (%)	104 (29)	132 (36)	0.04

Continued

Table 1 Continued

Characteristics	Male (n=367)	Female (n=370)	p Value
Patient reported depression questions			
Ever given clinical depression diagnosis	95 (26)	169 (46)	<0.01
Ever prescribed meds for depression	82 (22)	153 (41)	<0.01
Currently taking meds for depression	46 (13)	98 (26)	<0.01
Completed treatment for depression	1 (<0.1)	0	0.32
Treatment for depression successful	12 (3)	19 (5)	0.21
Ever stopped taking depression meds without telling clinician	21 (6)	37 (10)	0.03

*Not all column percentages sum to 100% due to missing values.
†Free care refers to a Massachusetts state programme for uninsured patients.
‡REALM health literacy categories correspond to total REALM scores of grade 3 and below (0–18), grades 4–6 (19–44), grades 7–8 (45–60), grade 9 and above (61–66).
§Prior hospital admissions and ED visits include those that occurred within 6 months prior to index admission.
¶Charlson comorbidity index reflects the cumulative increased likelihood of 1-year mortality; the higher the score, the more severe the burden of comorbidity; a 35% increase in risk of dying is reflected in a one-point increase in weights. Minimum score equals 0, no maximum score.
**Short Form-12 Health Survey (SF-12)—Physical Component Summary (PCS): range 0–100, mean score for US population=50 (SD=10), higher scores suggest greater physical functional status. Mental Component Summary (MCS): range 0–100, mean score for US population=50 (SD=10), higher scores suggest greater mental functional status.
††Patient Health Questionnaire (PHQ-9): 9-item 4-point Likert scale, standard scoring algorithm to diagnose major and minor depression and anxiety disorders.
ED, emergency department; PCP, primary care physician; REALM, Rapid Estimate of Adult Literacy in Medicine.

independent risk factor for returning to the hospital for men in this study.

Identifying and addressing risk factors associated with early post-discharge hospital utilisation is useful so that resources can be efficiently tailored to each individual patient's risk profile. Ideally, methods to ameliorate important risk factors are available. Some risk factors, like gender, however, may seem inherently immutable. Yet, as we demonstrated in this study, male gender is associated with other parameters that could potentially be effectively targeted.

Our findings raise the possibility that social isolation—as illustrated by the positive association with being retired, unmarried and symptoms of depression—may be important factors to target for intervention. Supporting these findings are studies examining

Table 2 Primary and secondary outcomes, by gender

	Men	Women	p Value
Primary outcomes ≤30 days after index hospitalisation			
Patients, n	367	370	
Hospital utilisations, n (visits/patient/mo)*	174 (0.474)	108 (0.292)	<0.01
IRR (95% CI)	1.62 (1.28 to 2.06)	REF	
Emergency department visits, n (visits/patient/mo)	101 (0.275)	50 (0.135)	<0.01
IRR (95% CI)	2.04 (1.45 to 2.86)	REF	
Readmissions, n (visits/patient/mo)	73 (0.199)	58 (0.157)	0.09
IRR (95% CI)	1.27 (0.90 to 1.79)	REF	
Secondary outcomes†			
Patients reached for 30-day follow-up call, n (%)	292 (80)	322 (87)	<0.01
Able to identify PCP name, n (%)	224 (77)	284 (88)	<0.01
PCP appt scheduled prior to discharge, n (%)	223 (60)	230 (63)	0.41
Visited PCP, n (%)	142 (49)	183 (57)	0.04
Visited specialist, n (%)	81 (28)	105 (33)	0.19
Able to identify discharge diagnosis, n (%)	212 (73)	247 (77)	0.24
How well did you understand your appointments after you left the hospital?‡ (those reporting understood well or very well)	210 (78%)	263 (87%)	<0.01
How well did you understand how to take your medications after leaving the hospital? (those reporting understood well or very well)	227 (84%)	270 (88%)	0.12
How well did you understand your main problem or diagnosis when you left the hospital? (those reporting understood well or very well)	175 (62%)	190 (61%)	0.65
How prepared were you to leave the hospital? (those reporting well prepared or very well)	175 (62%)	185 (59%)	0.40

*Defined as sum of emergency department (ED) visits plus rehospitalisations. Note: An ED visit that leads to a rehospitalisation is counted only as a rehospitalisation.
†Denominators reflect those subjects reached at 30-day follow-up phone call and those that answered question.
‡Questions asked on a 5-point Likert scale; Per cent reflects subjects who responded with either of the top two categories on the scale (ie, 'very prepared' or 'prepared').
IRR, incidence rate ratios.

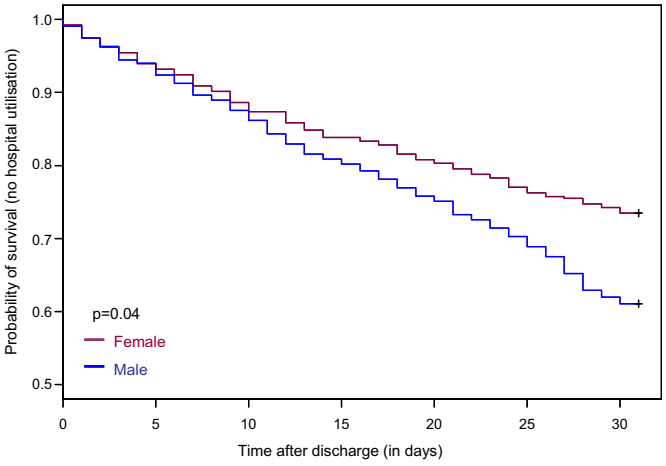


Figure 1 Kaplan–Meier curve: time to multiple hospitalisation events by gender.

the impact of social support and social networks. These studies have found that, in general, men are more socially isolated than women and that this contributes to worse health outcomes among men.[13 16] Men who were socially isolated were found to be less likely to undergo screening for blood pressure, cholesterol and cancer.[16]

Other studies suggest that men report less help-seeking behaviours, use primary care less[10] and are less likely to have a primary care physician when compared with women.[13] Overall, women use more health services than men due to pregnancy and cervical and breast cancer screening programmes.[6] However, lower rate of connectedness to primary care among men may also contribute to their excess use of hospital services and the finding that they may delay accessing care when it is needed.[17] Perhaps paradoxically, one study showed that increased access to primary care actually increased subsequent hospital utilisation; however, this study was conducted in the VA with almost exclusively male subjects and may have reflected appropriate use of hospital services among those who had been previously underserved.[18] Evidence suggests other factors that may impact a man's health-seeking behaviour including (1) men may have an overly optimistic perception of their health status, (2) the role women play in care-seeking decisions of men, (3) the influence of social networks and mood disorders and (4) the relatively lower value men appear to place on preventive care.[13 17]

Mood disorders can exacerbate the impact of social isolation on health. Men are far less likely than women to seek help for depression or anxiety.[19] Even when they do present for care, depression is often misdiagnosed or overlooked by providers.[20] These differences may be due to the differences in perceptions of distress experienced by men and women but may contribute to the low help-seeking behaviours exhibited by men. Still, given the effective treatment available for depression and anxiety disorders, depression represents a targetable risk for reducing unwarranted hospital utilisation by men.

Table 3 Adjusted incidence rate ratios (IRRs) and 95% CIs for socio-demographic characteristics on hospital utilisation in 30 days after index discharge

Variable	Total (n=737), IRR (95% CI)	Male (n=367), IRR (95% CI)	Female (n=370), IRR (95% CI)
Age	0.99 (0.98 to 1.00)	0.98 (0.97 to 0.99)	1.00 (0.99 to 1.02)
Gender			
Female	Ref	–	–
Male	1.55 (1.20 to 2.00)	–	–
Previous hospital visits (6 months)			
None	Ref	Ref	Ref
1–2	1.32 (0.95 to 1.83)	1.38 (0.89 to 2.13)	1.31 (0.78 to 2.21)
≥3	3.20 (2.35 to 4.35)	3.55 (2.38 to 5.29)	3.08 (1.86 to 5.10)
Employment status			
Employed	Ref	Ref	Ref
Not employed	1.13 (0.79 to 1.63)	1.31 (0.81 to 2.11)	0.97 (0.54 to 1.73)
Retired	2.29 (1.46 to 3.61)	3.27 (1.83 to 5.86)	1.10 (0.53 to 2.30)
Disabled	1.23 (0.86 to 1.75)	1.40 (0.88 to 2.23)	0.91 (0.53 to 1.59)
Other	1.53 (0.84 to 2.80)	0.86 (0.26 to 2.80)	1.70 (0.81 to 3.58)
Marital status			
Married	Ref	Ref	Ref
Not married	1.51 (1.10 to 2.06)	1.72 (1.12 to 2.64)	1.33 (0.83 to 2.15)
Positive depression screen	1.55 (1.20 to 2.00)	1.53 (1.09 to 2.13)	1.44 (0.96 to 2.15)
Study group, intervention	0.78 (0.61 to 1.00)	0.89 (0.65 to 1.23)	0.76 (0.50 to 1.13)
Visited PCP			
Yes	Ref	Ref	Ref
No	1.43 (1.07 to 1.93)	1.64 (1.08 to 2.50)	1.40 (0.92 to 2.14)
Not reached for follow-up call	2.16 (1.56 to 2.97)	2.19 (1.91 to 4.43)	1.07 (0.58 to 1.99)
Report PCP at baseline	0.96 (0.70 to 1.31)	1.22 (0.83 to 1.80)	0.57 (0.33 to 0.99)

IRR, incident rate ratio; PCP, primary care physician.

This analysis suggests that approaches to mitigate the risk of post-discharge rehospitalisations or ED visits among men may be to develop interventions that promote a connection to primary care, address social isolation and diagnose and treat depressive symptoms. Addressing these risks will require a creative and innovative approach including methods like routine screening for depressive symptoms, more aggressive empowerment of patients to engage the healthcare system proactively rather than reactively and establishing group visits within primary care to foster a social environment paired with the provision of primary care services and health education, as has been used in diabetes care and other chronic illnesses.[21]

This study has several limitations. Data on hospital utilisation outside Boston Medical Center were determined using patient self-report and not confirmed by EMR review at other hospitals. We were, however, able to confirm 91% of all events by consulting our own EMR. Second, our results may not be generalisable to populations other than those served by urban safety-net hospitals or other populations excluded from the RED trial (eg, non-English-speaking patients and patients admitted from nursing homes). Third, not all patients were reached at 30 days for the follow-up phone call, which is how the information regarding PCP follow-up was gathered. Finally, having done our project in Massachusetts, our population may have had an uncommonly high level of access to primary care.

In summary, our findings suggest that male gender is an important risk factor for early unplanned hospital utilisation within 30 days of discharge. This association may be linked to social behavioural patterns commonly associated with male gender, such as delayed help-seeking behaviours, often resulting in sporadic and episodic use of health services by men. Interventions targeting factors at the root of this phenomenon—such as social isolation, low rates of primary and preventive healthcare use and treatment of depressive symptoms—may help mitigate this gender effect. As health insurance reform and workforce development in primary care evolve, special efforts may be needed to acculturate men to the use of outpatient services.

Contributors SW was responsible for conception, design, acquisition of data, analysis and interpretation of results. She also made significant contributions to the drafting of the manuscript and revising for intellectual content and final approval of published version. SM was responsible for the analysis and interpretation of the results and made significant contributions to the drafting of the manuscript, its intellectual content and final approval of published version. CH was responsible for conception, design and acquisition of data and analysis and interpretation of the data. She contributed to the intellectual content of the manuscript and was responsible for approval of the final version of the article. MP was responsible for conception and design and analysis and interpretation of data. He also contributed significantly to the critical revisions of the manuscript and final approval of the submitted version. JG was responsible for conception and design and analysis and interpretation of data. He also contributed significantly to the critical revisions of the manuscript and final approval of the submitted version. VKC was responsible for conception and design, data analysis and interpretation, drafting of the manuscript and approval of the final version. JO was responsible for design, data collection, analysis and interpretation, drafting of the manuscript and final approval of published version. BJ was responsible for conception and design and analysis and interpretation of data. He also contributed significantly to the critical revisions of the manuscript and final approval of the submitted version.

Funding This project was supported by the Agency for Healthcare Research and Quality grants 1UC1HS014289-01 and 1U18HS015905-01 (BJ) and National Heart, Lung, and Blood Institute, National Institutes of Health, grant 1 R01 HL081307-01 (BJ).

Competing interests None.

Ethics approval Ethics approval was approved by the Boston University Internal Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No unpublished data are available.

REFERENCES

1. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare Fee-for-service program. *N Engl J Med* 2009;360:1418–28.
2. Patient Protection and Affordable Care Act, Pub. L. No. 111–148, 124 Stat. 119 (2010).
3. Oddone EZ, Weinberger M, Horner M, *et al*. Classifying general medicine readmissions. Are they preventable? Veterans affairs cooperative studies in health services group on primary care and hospital readmissions. *J Gen Intern Med* 1996;11:597–607.
4. Ashton CM, Wray NP. A conceptual framework for the study of early readmission as an indicator of quality of care. *Soc Sci Med* 1996;43:1533–41.
5. Jack BW, Chetty VK, Anthony D, *et al*. A Reengineered hospital discharge program to decrease rehospitalization. *Ann Intern Med* 2009;150:178–87.
6. Billings J, Dixon J, Mijanovich T, *et al*. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ* 2006;333:327.
7. Bottle A, Aylin P, Majeed A. Identifying patients at high risk of emergency hospital admissions: a logistic regression analysis. *J R Soc Med* 2006;99:406–14.
8. Weissman JS, Stern RS, Epstein AM. The impact of patient socioeconomic status and other social factors on readmission: a prospective study in four Massachusetts hospitals. *Inquiry* 1994;31:163–72.
9. Krumholz HM, Parent EM, Tu N, *et al*. Readmission after hospitalization for congestive heart Failure among medicare beneficiaries. *Arch Intern Med* 1997;157:99–104.
10. Cherpitel CJ. Emergency room and primary care services utilization and associated alcohol and drug use in the United States general population. *Alcohol Alcohol* 1999;34:581–9.
11. Mitchell SE, Paasche-Orlow MK, Forsythe SR, *et al*. Post-discharge hospital utilization among adult medical inpatients with depressive symptoms. *J Hosp Med* 2010;5:378–84.
12. Kartha A, Anthony D, Manasseh CS, *et al*. Depression is a risk factor for rehospitalization in medical inpatients. *Prim Care Companion J Clin Psychiatry* 2007;9:256–62.
13. Mackenzie CS, Gekoski WL, Knox VJ. Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. *Aging Ment Health* 2006;10:574–82.
14. Davis TC, Long SW, Jackson RH, *et al*. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med* 1993;25:391–5.
15. Charlson ME, Pompei P, Ales KL, *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
16. Molloy GJ, Perkins-Porras L, Strike PC, *et al*. Social networks and partner stress as predictors of adherence to medication, rehabilitation attendance, and quality of life following acute coronary syndrome. *Health Psychol* 2008;27:52–8.
17. Courtenay WH. Constructions of masculinity and their influence on men's well-being: a theory of gender and health. *Soc Sci Med* 2000;50:1385–401.
18. Weinberger M, Oddone EZ, Henderson WG. Does increased access to primary care Reduce hospital readmissions? *N Engl J Med* 1996;334:1441–7.
19. World Health Organization. *Gender and Mental Health*. 2002. http://whqlibdoc.who.int/gender/2002/a85573.pdf (accessed Jun 2011).
20. Kuehn BM. Men face barriers to mental health care. *JAMA* 2006;296:2303–4.
21. Davis AM, Sawyer DR, Vinci LM. The potential of group visits in diabetes care. *Clin Diabetes* 2008;26:58–62.

CHEST Supplement

ANTITHROMBOTIC THERAPY AND PREVENTION OF THROMBOSIS, 9TH ED: ACCP GUIDELINES

Antithrombotic Therapy for Atrial Fibrillation

Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines

John J. You, MD; Daniel E. Singer, MD; Patricia A. Howard, PharmD; Deirdre A. Lane, PhD; Mark H. Eckman, MD; Margaret C. Fang, MD, MPH; Elaine M. Hylek, MD, MPH; Sam Schulman, MD, PhD; Alan S. Go, MD; Michael Hughes, PhD; Frederick A. Spencer, MD; Warren J. Manning, MD; Jonathan L. Halperin, MD; and Gregory Y. H. Lip, MD

Background: **The risk of stroke varies considerably across different groups of patients with atrial fibrillation (AF). Antithrombotic prophylaxis for stroke is associated with an increased risk of bleeding. We provide recommendations for antithrombotic treatment based on net clinical benefit for patients with AF at varying levels of stroke risk and in a number of common clinical scenarios.**

Methods: **We used the methods described in the Methodology for the Development of Antithrombotic Therapy and Prevention of Thrombosis Guidelines: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines article of this supplement.**

Results: **For patients with nonrheumatic AF, including those with paroxysmal AF, who are (1) at low risk of stroke (eg, $CHADS_2$ [congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, prior stroke or transient ischemic attack] score of 0), we suggest no therapy rather than antithrombotic therapy, and for patients choosing antithrombotic therapy, we suggest aspirin rather than oral anticoagulation or combination therapy with aspirin and clopidogrel; (2) at intermediate risk of stroke (eg, $CHADS_2$ score of 1), we recommend oral anticoagulation rather than no therapy, and we suggest oral anticoagulation rather than aspirin or combination therapy with aspirin and clopidogrel; and (3) at high risk of stroke (eg, $CHADS_2$ score of ≥ 2), we recommend oral anticoagulation rather than no therapy, aspirin, or combination therapy with aspirin and clopidogrel. Where we recommend or suggest in favor of oral anticoagulation, we suggest dabigatran 150 mg bid rather than adjusted-dose vitamin K antagonist therapy.**

Conclusions: **Oral anticoagulation is the optimal choice of antithrombotic therapy for patients with AF at high risk of stroke ($CHADS_2$ score of ≥ 2). At lower levels of stroke risk, antithrombotic treatment decisions will require a more individualized approach.**

CHEST 2012; 141(2)(Suppl):e531S–e575S

Abbreviations: AAD = antiarrythmic drug; ACS = acute coronary syndrome; ACTIVE = Atrial Fibrillation Clopidogrel Trial With Irbesartan for Prevention of Vascular Events; AF = atrial fibrillation; AFASAK = Atrial Fibrillation Aspirin and Anticoagulation; AFFIRM = Atrial Fibrillation Follow-up Investigation of Rhythm Management; $CHADS_2$ = congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, prior stroke or transient ischemic attack; CHA_2DS_2-VASc = congestive heart failure (or left ventricular systolic dysfunction), hypertension, age ≥ 75 years, diabetes mellitus, prior stroke or transient ischemic attack, vascular disease, age 65 to 74 years, female sex; EAFT = European Atrial Fibrillation Trial; ESPS = European Stroke Prevention Study; INR = international normalized ratio; LAA = left atrial appendage; MI = myocardial infarction; PAF = paroxysmal atrial fibrillation; RCT = randomized controlled trial; RE-LY = Randomized Evaluation of Long-Term Anticoagulation Therapy; SPAF = Stroke Prevention in Atrial Fibrillation; TEE = transesophageal echocardiography; TIA = transient ischemic attack; VKA = vitamin K antagonist

Note on Shaded Text: Throughout this guideline, shading is used within the summary of recommendations sections to indicate recommendations that are newly added or have been changed since the publication of Antithrombotic and Thrombolytic Therapy: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Recommendations that remain unchanged are not shaded.

2.1.8. For patients with AF, including those with paroxysmal AF, who are at low risk of stroke (eg, $CHADS_2$ [congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, prior stroke or transient ischemic attack] score = 0), we suggest no therapy rather than antithrombotic therapy (Grade 2B)**. For patients who do choose antithrombotic therapy, we suggest aspirin (75 mg to 325 mg once daily) rather than oral anticoagulation** (Grade 2B) **or combination therapy with aspirin and clopidogrel** (Grade 2B)**.**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with antithrombotic therapy are likely to choose antithrombotic therapy rather than no antithrombotic therapy. Other factors that may influence the choices above are a consideration of patient-specific bleeding risk and the presence of additional risk factors for stroke, including age 65 to 74 years and female gender, which have been more consistently validated, and vascular disease, which has been less well validated (see section 2.1.12). The presence of multiple non-$CHADS_2$ risk factors for stroke may favor oral anticoagulation therapy.

2.1.9. For patients with AF, including those with paroxysmal AF, who are at intermediate risk of stroke (eg, $CHADS_2$ score = 1), we recommend oral anticoagulation rather than no therapy (Grade 1B)**. We suggest oral anticoagulation rather than aspirin (75 mg to 325 mg once daily)** (Grade 2B) **or combination therapy with aspirin and clopidogrel** (Grade 2B)**. For patients who are unsuitable for or choose not to take an oral anticoagulant (for reasons other than concerns about major bleeding), we suggest combination therapy with aspirin and clopidogrel rather than aspirin (75 mg to 325 mg once daily)** (Grade 2B)**.**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with anticoagulant therapy are likely to choose oral anticoagulation rather than antiplatelet therapy. Other factors that may influence the choice among antithrombotic therapies are a consideration of bleeding risk and the presence of additional risk factors for stroke, including age 65 to 74 years and female gender, which have been more consistently validated, and vascular disease, which has been less well validated (see section 2.1.12). The presence of multiple additional non-$CHADS_2$ risk factors for stroke may favor oral anticoagulation therapy.

2.1.10. For patients with AF, including those with paroxysmal AF, who are at high risk of stroke (eg, $CHADS_2$ score = 2), we recommend oral anticoagulation rather than no therapy (Grade 1A)**, aspirin (75 mg to 325 mg once daily)** (Grade 1B)**, or combination therapy with aspirin and clopidogrel** (Grade 1B)**. For patients who are unsuitable for or choose not to take an oral anticoagulant (for reasons other than concerns**

Revision accepted August 31, 2011.
Affiliations: From the Department of Medicine (Drs You, Schulman, and Spencer) and Department of Clinical Epidemiology and Biostatistics (Dr You), McMaster University, Hamilton, ON, Canada; Department of Medicine (Dr Singer), Harvard Medical School, and Clinical Epidemiology Unit (Dr Singer), General Medicine Division, Massachusetts General Hospital, Boston, MA; School of Pharmacy (Dr Howard), University of Kansas Medical Center, Kansas City, KS; University of Birmingham Centre for Cardiovascular Sciences (Drs Lane and Lip), City Hospital, Birmingham, England; Department of Clinical Medicine (Dr Eckman), Division of General Internal Medicine and Center for Clinical Effectiveness, University of Cincinnati, Cincinnati, OH; Department of Medicine (Dr Fang), Division of Hospital Medicine, University of California, San Francisco, San Francisco, CA; Boston University Medical Center Research Unit (Dr Hylek), Section of General Internal Medicine, Boston, MA; Comprehensive Clinical Research Unit (Dr Go), Division of Research, Kaiser Permanente Northern California, Oakland, CA; Decision Resources Inc (Dr Hughes), London, England; Section of Non-invasive Cardiac Imaging (Dr Manning), Beth Israel Deaconess Medical Center, Boston, MA; and The Cardiovascular Institute (Dr Halperin), Mount Sinai Medical Center, New York, NY.
Funding/Support: The Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines received support from the National Heart, Lung, and Blood Institute [R13 HL104758] and Bayer Schering Pharma AG. Support in the form of educational grants was also provided by Bristol-Myers Squibb; Pfizer, Inc; Canyon Pharmaceuticals; and sanofi-aventis US.

Correspondence to: Gregory Y. H. Lip, MD, University of Birmingham Centre for Cardiovascular Sciences, City Hospital, Birmingham, B18 7QH, England; e-mail: g.y.h.lip@bham.ac.uk

DOI: 10.1378/chest.11-2304

about major bleeding), we recommend combination therapy with aspirin and clopidogrel rather than aspirin (75 mg to 325 mg once daily) (Grade 1B)**.**

2.1.11. For patients with AF, including those with paroxysmal AF, for recommendations in favor of oral anticoagulation (including 2.1.9, 2.1.10, and excluding 2.2, 3.1, 3.2, 3.3), we suggest dabigatran 150 mg twice daily rather than adjusted-dose vitamin K antagonist (VKA) therapy (target INR range, 2.0-3.0) (Grade 2B)**.**

Remarks: Dabigatran is excreted primarily by the kidney. It has not been studied and is contraindicated in patients with severe renal impairment (estimated creatinine clearance of 30 mL/min or less). Clinicians should be aware that there is no antidote for dabigatran.

2.2. For patients with AF and mitral stenosis, we recommend adjusted-dose VKA therapy (target INR range, 2.0-3.0) rather than no therapy, aspirin (75 mg to 325 mg once daily), or combination therapy with aspirin and clopidogrel (all Grade 1B)**. For patients with AF and mitral stenosis who are unsuitable for or choose not to take adjusted-dose VKA therapy (for reasons other than concerns about major bleeding), we recommend combination therapy with aspirin and clopidogrel rather than aspirin (75 mg to 325 mg once daily) alone** (Grade 1B)**.**

3.1. For patients with AF and stable coronary artery disease (eg, no acute coronary syndrome within the previous year) and who choose oral anticoagulation, we suggest adjusted-dose VKA therapy alone (target international normalized ratio [INR] range, 2.0-3.0) rather than the combination of adjusted-dose VKA therapy and aspirin (Grade 2C)**.**

3.2. For patients with AF at high risk of stroke (eg, $CHADS_2$ score of 2 or greater) during the first month after placement of a bare-metal stent or the first 3 to 6 months after placement of a drug-eluting stent, we suggest triple therapy (eg, VKA therapy, aspirin, and clopidogrel) rather than dual antiplatelet therapy (eg, aspirin and clopidogrel) (Grade 2C)**. After this initial period of triple therapy, we suggest a VKA (INR 2.0-3.0) plus a single antiplatelet drug rather than VKA alone** (Grade 2C)**. At 12 months after intracoronary stent placement, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

For patients with AF at low to intermediate risk of stroke (eg, $CHADS_2$ score of 0 or 1) during the first 12 months after placement of an intracoronary stent (bare metal or drug eluting), we suggest dual antiplatelet therapy rather than triple therapy (Grade 2C)**. At 12 months after intracoronary stent placement, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with anticoagulant therapy are likely to choose triple therapy rather than dual antiplatelet therapy. Other factors that may influence this choice are a consideration of bleeding risk and the presence of additional non-$CHADS_2$ risk factors for stroke (see section 2.1.12).

3.3. For patients with AF at intermediate to high risk of stroke (eg, $CHADS_2$ score of 1 or greater) who experience an acute coronary syndrome and do not undergo intracoronary stent placement, we suggest for the first 12 months, adjusted-dose VKA therapy (INR 2.0-3.0) plus single antiplatelet therapy rather than dual antiplatelet therapy (eg, aspirin and clopidogrel) or triple therapy (eg, warfarin, aspirin, and clopidogrel) (Grade 2C)**. After the first 12 months, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

For patients with AF at low risk of stroke (eg, $CHADS_2$ score of 0), we suggest dual antiplatelet therapy (eg, aspirin and clopidogrel) rather than adjusted-dose VKA therapy (INR 2.0-3.0) plus single antiplatelet therapy or triple therapy (eg, warfarin, aspirin, and clopidogrel) (Grade 2C)**. After the first 12 months, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with anticoagulant therapy are likely to choose adjusted-dose VKA therapy plus single antiplatelet therapy rather than dual antiplatelet therapy. Other factors that may influence this choice are a consideration of bleeding risk and the presence of additional non-$CHADS_2$ risk factors for stroke (see section 2.1.12).

3.4. For patients with AF being managed with a rhythm control strategy (pharmacologic or

catheter ablation), we suggest that antithrombotic therapy decisions follow the general risk-based recommendations for patients with AF in section 2.1, regardless of the apparent persistence of normal sinus rhythm (Grade 2C)**.**

3.5. For patients with atrial flutter, we suggest that antithrombotic therapy decisions follow the same risk-based recommendations as for AF.

4.1.1. For patients with AF of greater than 48 h or unknown duration undergoing elective electrical or pharmacologic cardioversion, we recommend therapeutic anticoagulation (adjusted-dose VKA therapy, target INR range 2.0-3.0, low-molecular-weight heparin at full venous thromboembolism treatment doses, or dabigatran) for at least 3 weeks before cardioversion or a transesophageal echocardiography (TEE)-guided approach with abbreviated anticoagulation before cardioversion rather than no anticoagulation (Grade 1B)**. We recommend therapeutic anticoagulation for at least 4 weeks after successful cardioversion to sinus rhythm rather than no anticoagulation, regardless of the baseline risk of stroke** (Grade 1B)**. Decisions about anticoagulation beyond 4 weeks should be made in accordance with our risk-based recommendations for long-term antithrombotic therapy in section 2.1.**

4.1.2. For patients with AF of documented duration of 48 h or less undergoing elective cardioversion (electrical or pharmacologic), we suggest starting anticoagulation at presentation (low-molecular-weight heparin or unfractionated heparin at full venous thromboembolism treatment doses) and proceeding to cardioversion rather than delaying cardioversion for 3 weeks of therapeutic anticoagulation or a TEE-guided approach (Grade 2C)**. After successful cardioversion to sinus rhythm, we recommend therapeutic anticoagulation for at least 4 weeks rather than no anticoagulation, regardless of baseline stroke risk** (Grade 2C)**. Decisions about long-term anticoagulation after cardioversion should be made in accordance with our risk-based recommendations for long-term antithrombotic therapy in section 2.1.**

4.2. For patients with AF and hemodynamic instability undergoing urgent cardioversion (electrical or pharmacologic), we suggest that therapeutic-dose parenteral anticoagulation be started before cardioversion, if possible (Grade 2C)**, but that initiation of anticoagulation must not delay any emergency intervention** (Grade 2C)**. After successful cardioversion to sinus rhythm, we suggest therapeutic anticoagulation for at least 4 weeks after successful cardioversion to sinus rhythm rather than no anticoagulation, regardless of baseline stroke risk** (Grade 2C)**. Decisions about anticoagulation beyond 4 weeks should be made in accordance with our risk-based recommendations for long-term antithrombotic therapy in section 2.1.**

4.3. For patients with atrial flutter undergoing elective or urgent pharmacologic or electrical cardioversion, we suggest that the same approach to thromboprophylaxis be used as for patients with atrial fibrillation undergoing cardioversion.

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia. One in four individuals aged 40 years will develop AF during his or her lifetime, and it has been estimated that by the year 2050, up to 16 million Americans will have AF.[1,2] Nonrheumatic AF is a strong, independent predictor of ischemic stroke associated with a fivefold increase in risk.[3] Without thromboprophylaxis, the risk of ischemic stroke in patients with nonrheumatic AF, as seen in the control arms of the original trials of antithrombotic therapy in AF, is ~5% per year.[4] Over the past 2 decades, considerable work has been done to evaluate antithrombotic therapies to prevent stroke in patients with AF, and the field continues to evolve with the emergence of a new generation of oral anticoagulants.

This article begins with a discussion of the methods used to develop our recommendations for antithrombotic therapy in patients with AF. Next, we provide our treatment recommendations, divided into the following sections:

1. Antithrombotic therapy in patients with AF in general (includes patients with permanent, persistent, or paroxysmal AF [PAF])
2. Antithrombotic therapy in patients with AF in special situations:
   - Stable coronary artery disease
   - Acute coronary syndrome (ACS)
   - Intracoronary artery stent
   - Acute ischemic stroke
   - Management with a rhythm control strategy
   - Chronic atrial flutter
3. Antithrombotic therapy for patients with AF undergoing cardioversion

The article ends with a discussion of practical issues in the use of adjusted-dose vitamin K antagonist (VKA) therapy in patients with AF and suggestions for future research.

Table 1 specifies the clinical question being addressed in this article (in PICO [population, intervention, comparator, outcomes] format) and the types of studies used. This article does not give recommendations for antithrombotic therapy in patients with AF around the time of surgical or invasive procedures (see Douketis et al[5]), at the time of presentation with acute stroke (see Lansberg et al[6]), or in patients with AF who have prosthetic heart valves (see Whitlock et al[7]). This article does not give recommendations for patients with AF who are pregnant. For general recommendations on antithrombotic therapy during pregnancy (ie, not specific to AF), see Bates et al.[8] Finally, the recommendations in this article apply to patients with persistent and permanent AF and to patients with PAF but do not apply to patients with a single, transient, self-limited episode of AF associated with acute illness.

1.0 Methods

To inform our guideline development, we searched for relevant articles published since the last literature search performed for the Antithrombotic and Thrombolytic Therapy: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Specifically, for literature regarding the assessment of stroke risk in patients with AF, we searched Medline for articles published from January 1, 2005, to October 2009 using the search terms "atrial fibrillation," "atrial flutter," "risk assessment," "risk factors," "risk stratification," "stroke," and "thromboembolism." For literature regarding prevention of stroke and thromboembolism in patients with AF, we searched Medline for articles published from January 1, 2005, to October 2009 using the search terms "coumarins," "warfarin," "dicumarol," "phenprocoumon," "acenocoumarol," "fondaparinux," "idraparinux," "aspirin," "triflusal," "indobufen," "dabigatran," "ximelagatran," "rivaroxaban," "apixaban," "ticlopidine," "clopidogrel," "catheter ablation," "watchman," "PLAATO," "cardioversion," "atrial fibrillation," and "atrial flutter."

1.1 Outcomes of Interest

The outcomes most relevant to patients with AF include death, nonfatal stroke, systemic embolism, nonfatal major extracranial bleeding, and the burden and lifestyle limitations associated with outpatient antithrombotic therapy. To facilitate decision-making, the term stroke in this article includes ischemic stroke and intracranial hemorrhage (intracerebral, subdural, and subarachnoid hemorrhage). Although there may be some differences in the impact of these intracranial events (eg, subdural hemorrhage) on quality of life, we judged that on average, the impact would be similar. We also explicitly considered that outpatient antithrombotic therapy was associated with a burden to the patient, which in some cases, such as aspirin, is a requirement to take a daily medication, or in other cases, such as adjusted-dose VKA therapy, is not only a requirement to take a daily medication but also a requirement to limit one's lifestyle, restrict one's diet, and undergo frequent blood testing and clinic visits. For recommendations about patients with AF and stable coronary artery disease, intracoronary stent placement, or recent ACS (sections 3.1-3.3), we also considered the effect of different treatment options on the outcome of nonfatal myocardial infarction (MI).

1.2 Patient Values and Preferences

In developing our treatment recommendations, we attempted to account for patient values and preferences regarding these health states. In this guideline, a systematic review of studies assessing values and preferences related to antithrombotic therapy found that values for health states and preferences for treatments vary appreciably among individuals (MacLean et al[9]). The available literature has several limitations. The studies eliciting preferences were small and used different methods and tools, and most included a sizeable proportion of participants who had previously taken or were currently taking VKAs. Furthermore, there are inconsistencies across studies that often are difficult to explain, leaving considerable uncertainty about average patient values. Nevertheless, to make our recommendations, we required an estimate of average patient values for the relevant outcomes so that we could judge whether the trade-offs between benefit and harm would favor one course of treatment over another.

As described by Guyatt et al[10] in this guideline, to obtain our estimates of average patient values, we have used ratings of key health states from participating guideline panelists informed by our systematic review of the relevant literature. The results of the panelist value rating exercise suggest that on average, patients would find a typical nonfatal stroke (ischemic or hemorrhagic) approximately three times as aversive as a nonfatal major extracranial bleed (typically a GI bleed) and a typical nonfatal MI as aversive as a nonfatal major extracranial bleed.

With the exception of the choice between VKA therapy and no therapy, the choice of one antithrombotic treatment over another for long-term stroke prevention in AF will not lead to differences in all-cause mortality (section 2). Thus, for these choices that are not expected to result in a difference in mortality, for every 1,000 patients treated for 1 year, if the number of nonfatal strokes prevented is less than one-third of the number of nonfatal major extracranial bleeds caused by a given antithrombotic therapy, we have recommended against that intervention. If the number of nonfatal strokes prevented is appreciably more than one-third of the number of nonfatal major extracranial bleeding events that result from a given antithrombotic therapy, we have recommended in favor.

Making these trade-offs requires not only estimates regarding average patient values for the relevant outcomes but also best estimates of (1) the effect of a given treatment on these outcomes against a given comparator (ie, relative risk) and (2) the absolute event rates for these outcomes in untreated patients (or, for PICO questions of a treatment vs an active comparator, the absolute event rates in patients receiving the active comparator). In the following section, we present the methods used to obtain these estimates for our guidelines.

1.3 Estimating the Magnitude of Treatment Effect

For each clinical question, we extracted data regarding the previously discussed outcomes from the relevant clinical trials. When there were multiple randomized controlled trials (RCTs) addressing the same clinical question, we conducted meta-analyses using random-effects models and the Mantel-Haenszel method to obtain pooled estimates of treatment effect, expressed as relative risk. For studies that did not report the proportion of strokes that were fatal and nonfatal, we used all available data in the published report to obtain a best estimate of the effect of

Table 1—*[Introduction] Clinical Questions Addressed in This Article*

Section	Population	Intervention	Comparator	Outcomes of Interest	Available Methodology
		2. Patients with AF in general			
2.1.1	Patients with nonrheumatic AF	Adjusted-dose VKA	No therapy	Death	RCTs
2.1.2		Antiplatelet monotherapy (eg, aspirin)	No therapy	Nonfatal strokes Nonfatal major extracranial bleeds	
2.1.3		VKA	Antiplatelet monotherapy (eg, aspirin)	Systemic embolism Procedural complications (for percutaneous closure of left atrial appendage only)	
2.1.4		VKA	Aspirin + clopidogrel		
2.1.5		Aspirin + clopidogrel	Aspirin		
2.1.6		New oral anticoagulants	VKA		
2.1.11		Dabigatran	VKA		
2.1.13		Percutaneous closure of left atrial appendage	VKA		RCTs Cohort studies
2.2	Patients with AF and mitral stenosis	VKA	No therapy, antiplatelet monotherapy (eg, aspirin), or aspirin + clopidogrel		Cohort studies
		3. Management of antithrombotic therapy for patients with AF in special situations			
3.1	Patients with AF and stable coronary artery disease	VKA + aspirin	VKA	Death Nonfatal strokes	Cohort studies
3.2	Patients with AF and placement of an intracoronary stent (with or without recent ACS)	VKA + aspirin + clopidogrel	Aspirin + clopidogrel	Nonfatal MI Nonfatal major extracranial bleeds Systemic embolism	
3.3	Patients with AF and ACS who do not undergo intracoronary stent placement	VKA + aspirin	Aspirin + clopidogrel VKA + aspirin + clopidogrel		
3.4	Patients with AF being managed with a rhythm control strategy	VKA	No VKA	Death Nonfatal strokes	
3.5	Patients with chronic atrial flutter	Antithrombotic therapy options as per section 2.1		Nonfatal major extracranial bleeds Systemic embolism	
		4. Patients with AF undergoing cardioversion			
4.1.1	Patients with AF of > 48 h or unknown duration undergoing elective cardioversion	Minimum 3 wk anticoagulation before and 4 wk after cardioversion	No anticoagulation	Death Nonfatal strokes Nonfatal major extracranial bleeds	Cohort studies
		Abbreviated precardioversion anticoagulation + TEE-guided cardioversion	No anticoagulation	Systemic embolism	RCTs Cohort studies
4.1.2	Patients with AF of ≤ 48 h duration undergoing elective cardioversion	Anticoagulation before cardioversion	No anticoagulation before cardioversion		Cohort studies
4.2	Patients undergoing urgent cardioversion for hemodynamically unstable AF	Anticoagulation before cardioversion	No anticoagulation before cardioversion		Cohort studies
4.3	Patients undergoing elective or urgent cardioversion for atrial flutter	Anticoagulation before and after cardioversion	No anticoagulation		Cohort studies

ACS = acute coronary syndrome; AF = atrial fibrillation; MI = myocardial infarction; RCT = randomized controlled trial; TEE = transesophageal echocardiography; VKA = vitamin K antagonist.

treatment on nonfatal stroke by assuming a case fatality rate of 50% for hemorrhagic stroke and 25% for ischemic stroke based on population-based stroke registry data and the case fatality rates observed in the RCTs of patients with AF that reported such data.[11] For nonfatal major extracranial bleeding, we accepted the definition of major bleeding from the individual studies, and when the proportion of major extracranial bleeds that were fatal vs nonfatal was not reported, we applied the average case fatality rate for major extracranial bleeding reported across the relevant clinical trials (~15%). For the outcome of systemic embolism, we used the total number of events (fatal and nonfatal) because systemic embolism was an infrequent event and typically not reported as fatal vs nonfatal.

1.4 Deriving Baseline Risk of Stroke in Patients With AF

The risk of stroke varies considerably across different groups of patients with AF.[12] Simple validated tools for the assessment of stroke risk in patients with AF help in identifying those who are more likely to benefit than be harmed from antithrombotic therapy.

1.4.1 Pattern of AF: Guidelines have categorized the pattern of AF into (1) PAF, in which recurrent episodes terminate spontaneously within 7 days and usually in < 48 h; (2) persistent AF, in which the episode of AF does not self-terminate within 7 days or is terminated by cardioversion; and (3) permanent AF, in which AF is present for some time and cardioversion either has failed or has not been attempted.[13]

For patients with PAF, periods of sinus rhythm theoretically should lessen the risk of stroke, yet transitions from AF to sinus rhythm may acutely heighten risk in a manner similar to the increase in risk caused by cardioversion. Although some studies suggest that PAF is associated with a lower risk of stroke than persistent or permanent AF, patients with PAF generally are younger and have a lower prevalence of other stroke risk factors, and clinical trial data suggest that PAF confers a relative risk of stroke similar to persistent or permanent AF when controlling for associated stroke risk factors.[14-17] Pending further evidence, it seems reasonable to treat patients with PAF in a manner similar to those with persistent and permanent AF; thus, our risk-based treatment recommendations apply to patients with PAF and persistent and permanent AF.

1.4.2 Independent Risk Factors for Stroke in Patients With AF: Two recent systematic reviews have identified clinical and echocardiographic factors that are independently associated with an increased risk of stroke in patients with AF.[18,19] The individual studies from these systematic reviews, in addition to articles identified in an updated literature search performed for this guideline, are summarized in Tables S1 through S12 (tables that contain an "S" before the number denote supplementary tables not contained in the body of the article and available instead in an online data supplement; see the "Acknowledgments" for more information). In terms of clinical applicability, the most consistently identified risk factors for ischemic stroke among patients with AF are a history of ischemic stroke or transient ischemic attack (TIA)—the strongest dichotomous predictor of stroke risk—older age, hypertension, and diabetes. Although impaired left ventricular systolic function is a risk factor for stroke in AF, there are conflicting data about whether a history of congestive heart failure per se raises the risk of ischemic stroke in AF. Although age thresholds often are used in stroke risk schemes, stroke risk in AF increases continuously with age, appreciably rising from age 65 years onward.[20] There is moderate-quality evidence that women face a higher risk of stroke than men.[18] A history of coronary artery disease has not consistently been found to be an independent risk factor for stroke in patients with AF.[18] However, there is low-quality evidence that the presence of atherosclerotic vascular disease (eg, complex aortic plaque in the descending aorta seen on transesophageal echocardiography [TEE] or a history of peripheral arterial disease) independently predicts stroke risk among patients with AF.[21-23]

1.4.3 Stroke Risk Stratification Schema: Many risk stratification schema, which use various combinations of the risk factors discussed previously, have been developed to aid clinicians in the assessment of stroke risk in patients with nonrheumatic AF (Table S13).[12] Despite substantial efforts in this field over the past several decades, all available schema have only modest ability to predict stroke in patients with AF, with C statistics typically between 0.55 and 0.70.[4,24-43] (A C statistic of 0.50 indicates a model that does not discriminate better than chance alone, and a C statistic of 1.00 indicates perfect discrimination.)

The $CHADS_2$ score is the most validated risk scheme, having been independently tested in at least 10 separate cohorts after its original derivation.[24,32,35-46] The $CHADS_2$ score gives a single point for each of congestive heart failure (originally defined as a recent exacerbation of congestive heart failure), hypertension (defined as a history of hypertension, rather than a presence of elevated BP), age ≥ 75 years, and diabetes mellitus and two points for prior stroke or TIA (Table 2).[30]

Despite its widespread adoption and ease of use, the $CHADS_2$ score has limitations. First, congestive heart failure is not a consistently demonstrated independent predictor of stroke. Second, the risk associated with a history of hypertension may differ among patients with well-treated vs poorly treated hypertension.[46-48] Third, in most studies, the $CHADS_2$ score has only a modest ability to predict stroke in patients with AF (C statistic, 0.56-0.70).[24,32,35-46] Finally, the threshold of stroke risk at which treatment with oral anticoagulation will be preferred is likely to decrease with the emergence of new oral anticoagulants that do not require regular monitoring of the international normalized ratio (INR) and that may be associated with greater reductions in stroke and less risk of bleeding compared with adjusted-dose VKA therapy.[49] Thus, stroke risk stratification schema will need to evolve to more accurately identify patients who are at sufficiently low risk of stroke and can be treated with aspirin or no antithrombotic therapy, whereas all other patients with AF can be considered for oral anticoagulation.

The CHA_2DS_2-VASc (congestive heart failure [or left ventricular systolic dysfunction], hypertension, age < 75 years; diabetes mellitus, prior stroke or transient ischemic attack, vascular disease, age 65 to 74 years, female sex) score is a new risk scheme that combines the $CHADS_2$ score with additional moderate risk factors, which were also included in the 2006 American College of Cardiology/American Heart Association/European Society of Cardiology and National Institute for Health and Clinical Excellence AF practice guidelines.[29,50] Specifically, the CHA_2DS_2-VASc score assigns points as in the original $CHADS_2$ score (Table 2) with the exception of age ≥ 75 years, which is assigned two points. It also assigns a single point for each of the following additional

Table 2—*[Section 1.4.3] $CHADS_2$ Score[30] for Assessment of Stroke Risk in Patients With Nonrheumatic AF*

Risk Factor	Points
Recent **C**ongestive heart failure exacerbation	1
History of **H**ypertension	1
Age ≥ 75 y	1
Diabetes mellitus	1
Prior history of **S**troke or transient ischemic attack	2

$CHADS_2$ = congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, prior stroke or transient ischemic attack. See Table 1 legend for expansion of other abbreviation.

risk factors: female sex, age 65 to 74 years, and vascular disease (defined as a history of MI, peripheral arterial disease, or complex aortic plaque). The CHA_2DS_2-VASc score has been evaluated in at least five separate cohorts since its original description. With the exception of a recent study by Olesen et al,[24] all other studies have found that the predictive ability of CHA_2DS_2-VASc is similar to that of the $CHADS_2$ score (C statistics of each risk score is ~0.6 across the various studies) and not statistically significantly greater than that of $CHADS_2$.[36,40-43] Because the $CHADS_2$ score has been extensively validated and is easy for clinicians to remember and use, we use the $CHADS_2$ score as the principal approach for our risk-based treatment recommendations.

1.4.4 Estimating the Baseline Risk of Nonfatal Stroke by $CHADS_2$ Score: To develop our recommendations, we required estimates of the absolute rate of nonfatal stroke (ischemic or hemorrhagic) for patients according to their underlying risk of stroke, as characterized by their $CHADS_2$ score. Ideally, we would obtain these estimates of baseline risk from published annual rates of stroke, by $CHADS_2$ score, among untreated patients. However, such data are not available.

Therefore, to obtain estimates of annual stroke risk, we used pooled data from aspirin-treated patients enrolled in six clinical trials of antithrombotic therapy for stroke prevention in AF.[37] This published report presented data regarding ischemic stroke rates (fatal and nonfatal combined) on aspirin, stratified by $CHADS_2$ score. We used the following calculations to estimate the annual risk of nonfatal stroke (ie, ischemic and hemorrhagic) on aspirin: (1) multiplication of reported rates of ischemic stroke by 1.08 to account for additional hemorrhagic strokes on aspirin therapy (based on the observed ratio of ischemic:hemorrhagic strokes in aspirin arms of RCTs in patients with AF) and (2) an estimation that 50% and 25% of hemorrhagic and ischemic strokes, respectively, were fatal.

Depending on the clinical question being addressed by a particular recommendation, we adjusted these absolute rates of nonfatal stroke on aspirin to reflect the clinical scenario being addressed. For the recommendation addressing VKA therapy vs no treatment, for example, to estimate absolute rates of nonfatal stroke on no treatment, we increased our estimates of the rate of nonfatal stroke on aspirin by 21% to account for the estimated efficacy of aspirin in preventing stroke in AF). We therefore used the following absolute rates of nonfatal stroke in untreated patients to develop our recommendations: 0.8%, 2.2%, 4.5%, and 9.6% per year for patients with $CHADS_2$ scores of 0, 1, 2, and 3 to 6, respectively.

We have chosen to base our treatment recommendations on absolute rates of stroke derived from clinical trials, recognizing that these data have important limitations. Less than 10% of patients screened were enrolled in these historical trials, there was limited racial and ethnic diversity, and there is some evidence suggesting that stroke rates may now be lower than at the time these RCTs were conducted 2 decades ago possibly because of improved treatment of cardiovascular risk factors, such as hypertension.[51,52] However, despite these limitations, these clinical trial-based data regarding stroke events were systematically and prospectively collected, and they remain the best available source of stroke rates stratified by $CHADS_2$ score.

1.5 Deriving Baseline Risk of Death in Patients With AF

We used data from an observational health plan database study of 11,526 patients with nonvalvular AF (the Anticoagulation and Risk Factors in Atrial Fibrillation [ATRIA] Study) to obtain an estimate of the risk of all-cause mortality in patients with AF not treated with warfarin (53 deaths per 1,000 patient-years). Untreated patients in this cohort had a mean $CHADS_2$ score of 1.5, and a substantial majority (78%) of patients in the cohort had a $CHADS_2$ score of ≥ 1.[44] Although $CHADS_2$-specific rates of all-cause mortality are not available from this cohort, the risk of death is expected to be lower in low-risk patients with a $CHADS_2$ score of 0 because of their younger age, lower prevalence of vascular risk factors, and lower rates of fatal ischemic stroke.

1.6 Deriving Baseline Risk of Nonfatal Major Extracranial Bleeding

To develop our recommendations, we also required estimates of the baseline risk of nonfatal major extracranial bleeding in patients with AF. We obtained this estimate from observational studies of VKA therapy in cohorts that included exclusively or predominantly patients with AF (median rate of 1.3% per year across these studies).[44,53,54] To estimate the baseline risk of nonfatal major extracranial bleeding off VKA therapy, we used this median rate of bleeding on VKA therapy from the observational studies (1.3% per year) and divided by the pooled relative risk (2.58) of nonfatal major extracranial bleeding associated with VKA therapy, as obtained from RCTs of VKA therapy vs no therapy. Therefore, our estimate of the baseline risk of nonfatal major extracranial bleeding off therapy was 0.5% per year.

2.0 Antithrombotic Therapy for Patients With AF in General

Over the past 2 decades, numerous RCTs have investigated antithrombotic therapies to reduce the risk of thromboembolism, principally ischemic stroke, in patients with AF. In this section, we summarize the evidence and give treatment recommendations for VKA therapy, antiplatelet monotherapy (eg, aspirin), dual antiplatelet therapy with aspirin and clopidogrel, and new oral anticoagulants (eg, dabigatran) in patients with AF.

2.1 Patients With Nonrheumatic AF

2.1.1 VKAs vs No Therapy: Six RCTs that enrolled a total of 2,584 patients and address the primary prevention (Atrial Fibrillation Aspirin and Anticoagulation [AFASAK] 1, Boston Area Anticoagulation Trial for Atrial Fibrillation [BAATAF], Canadian Atrial Fibrillation Anticoagulation [CAFA], Stroke Prevention in Atrial Fibrillation [SPAF] I, Stroke Prevention in Nonrheumatic Atrial Fibrillation [SPINAF]) and secondary prevention (European Atrial Fibrillation Trial [EAFT]) of stroke in patients with AF provide high-quality evidence that VKA therapy reduces the risk of death by one-fourth and the risk of nonfatal stroke by two-thirds compared with no therapy (Table 3).[55-61] For patients with a $CHADS_2$ score of 0, the studies provide moderate-quality evidence that VKA therapy increases the risk of nonfatal major extracranial bleeding due to imprecision of the estimate (wide CIs). The studies provide high-quality evidence for patients with higher $CHADS_2$ scores.

Table 3—*[Sections 2.1.1, 2.1.8, 2.1.9, 2.1.10] Should VKAs Rather Than No Therapy Be Used in Patients With AF?*[a]

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1 y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With No Therapy	With VKA	Relative Effect (95% CI)	With No Therapy	With VKA (95% CI)	Quality of Evidence
Death (critical outcome) mean follow-up 1.7 y											
6 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	136/1,425 (9.5)	103/1,429 (7.2)	RR 0.72 (0.55-0.94)	53 per 1,000[b]	15 fewer deaths per 1,000[c] (from 3 fewer to 24 fewer)	High
Nonfatal stroke (critical outcome) mean follow-up 1.7 y; ischemic stroke and intracranial hemorrhage[d]											
6 RCTs	No serious limitations	No serious inconsistency	No serious Indirectness	No serious imprecision	Undetected	108/1,425[e] (7.6)	36/ 1,429[f] (2.5)	RR 0.34 (0.23-0.49)	$CHADS_2$ 0 points		High
									8 per 1,000[g]	5 fewer strokes per 1,000 (from 4 fewer to 6 fewer)	
									$CHADS_2$ 1 point		
									22 per 1,000	15 fewer strokes per 1,000 (from 11 fewer to 17 fewer)	
									$CHADS_2$ 2 points		
									45 per 1,000	30 fewer strokes per 1,000 (from 23 fewer to 35 fewer)	
									$CHADS_2$ 3-6 points		
									96 per 1,000	63 fewer strokes per 1,000 (from 49 fewer to 74 fewer)	
Nonfatal major extracranial bleeds (important outcome) mean follow-up 1.7 y											
6 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	7/1,425 (0.5)	23/1,429 (1.6)	RR 2.58 (1.12-5.97)	5 per 1,000[h]	8 more bleeds per 1,000 (from 1 more to 25 more)	High[i]
Systemic embolism (important outcome) mean follow-up 1.7 y											
6 RCTs	No serious limitations	No serious inconsistency	No serious Indirectness	Imprecise[j]	Undetected	10/1,425 (0.7)	4/1,429 (0.3)	RR 0.42 (0.15-1.20)	4 per 1,000[k]	2 fewer systemic emboli per 1,000 (from 3 fewer to 1 more)	Moderate

(Continued)

Table 3—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1 y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With No Therapy	With VKA	Relative Effect (95% CI)	With No Therapy	With VKA (95% CI)	Quality of Evidence
Burden of treatment (important outcome)											
N/A	No serious limitations	No serious limitations	No serious limitations	No serious limitations	N/A	N/A	N/A	VKA > no therapy	Warfarin: daily medication, lifestyle limitations, dietary restrictions, frequent blood testing and clinic visits		High

AFASAK = Atrial Fibrillation Aspirin and Anticoagulation; ATRIA = Anticoagulation and Risk Factors in Atrial Fibrillation; BAATAF = Boston Area Anticoagulation Trial for Atrial Fibrillation; CAFA = Canadian Atrial Fibrillation Anticoagulation; EAFT = European Atrial Fibrillation Trial; IPD = individual patient data; N/A = not applicable; RR = risk ratio; SPAF = Stroke Prevention in Atrial Fibrillation; SPINAF = Stroke Prevention in Nonrheumatic Atrial Fibrillation. See Table 1 and 2 legends for expansion of other abbreviations.

[a]Pooled estimates of treatment effect in this evidence profile are from a meta-analysis conducted for these guidelines, including data from 6 RCTs of adjusted-dose VKA therapy vs no antithrombotic therapy (AFASAK I, BAATAF, CAFA, EAFT, SPAF I, SPINAF).

[b]From Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort). The majority (78%) of the patients in ATRIA had a $CHADS_2$ score of ≥ 1.

[c]VKA therapy likely does not lead to any reductions in all-cause mortality compared with no therapy in low-risk patients with a $CHADS_2$ score of 0 because there is evidence suggesting that absolute reductions in ischemic stroke heavily depend on stroke risk (eg, small reductions at low $CHADS_2$ score), whereas the absolute increase in intracranial hemorrhage is relatively consistent across $CHADS_2$ categories (Singer et al[51]).

[d]Intracranial hemorrhage includes intracerebral, subdural, and subarachnoid bleeds. For studies that did not report the number of strokes that were fatal and nonfatal (CAFA, EAFT), we imputed values for the number of fatal and nonfatal strokes to estimate the pooled relative risk for nonfatal stroke across all eligible studies. Assumptions underlying these estimates are detailed in the section 1.3 of this article.

[e]Of the 108 nonfatal strokes on no therapy, an estimated 106 (98%) were ischemic, and two (2%) were hemorrhagic.

[f]Of the 36 nonfatal strokes on VKA therapy, an estimated 34 (94%) were ischemic, and two (6%) were hemorrhagic.

[g]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on no therapy by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in the section 1.4.4 of this article.

[h]Estimate of the rate of nonfatal major extracranial bleeding on no therapy is derived from rates in observational cohorts of predominantly patients with AF receiving adjusted-dose VKA therapy (median, 1.3%/y) and dividing by the relative risk of major bleeding on adjusted-dose VKA therapy observed in the RCTs (see section 1.6 of this article).

[i]Quality of evidence is moderate for patients with a $CHADS_2$ score of 0 because the recommendation would differ if the true increase in nonfatal major extracranial bleeding was one vs 25 bleeds per 1,000 patients in a given year.

[j]The 95% CI does not exclude the possibility of no effect.

[k]Based on rate of systemic embolism on aspirin of 0.3 per 100 patient-y reported in the IPD meta-analysis of warfarin vs aspirin by van Walraven et al[61] and a relative risk of 0.80 for aspirin vs no therapy for systemic embolism.

The studies also provide moderate-quality evidence that VKA therapy reduces the risk of systemic embolism (rated down for imprecision).

2.1.2 Antiplatelet Monotherapy (Aspirin) vs No Therapy: Several RCTs of antiplatelet monotherapy vs no therapy in patients with AF have shown that antiplatelet therapy leads to, at best, a modest reduction in the risk of nonfatal stroke. Antiplatelet monotherapy in these trials was either with aspirin alone (AFASAK 1; SPAF I; EAFT; United Kingdom Transient Ischaemic Attack Aspirin trial [UK-TIA]; Low-Dose Aspirin, Stroke, Atrial Fibrillation [LASAF]; European Stroke Prevention Study [ESPS]-2; and Japan Atrial Fibrillation Stroke Trial [JAST]), aspirin in combination with fixed (ineffective) minidose warfarin (Swedish Atrial Fibrillation Trial [SAFT]), or dipyridamole (ESPS-2).[55,58,60,62-66] Aspirin dosing in these trials typically ranged from 50 to 325 mg/d. In the LASAF, JAST, and SAFT trials, aspirin was compared with a no-treatment control arm, whereas in the remaining trials, comparison was to placebo.[63,64,66] ESPS-2 and UK-TIA were stroke prevention trials conducted primarily in non-AF populations, and only the data from the subset of patients with AF are considered here.[62,65]

Pooled data from these trials provide moderate-quality evidence (rated down for imprecision) that antiplatelet monotherapy is associated with a 21% relative reduction in risk of nonfatal stroke compared with no treatment (Table 4). Although our confidence in the benefits of aspirin therapy is moderate, our confidence in its bleeding risk is high. Although the trials of antiplatelet monotherapy in patients with AF were underpowered to precisely estimate the risk of nonfatal major extracranial bleeding, trials of aspirin for the primary and secondary prevention of cardiovascular disease have conclusively demonstrated that aspirin is associated with an increased risk of major hemorrhage. An individual patient data meta-analysis combining data from six cardiovascular primary prevention trials (95,000 subjects) and a meta-analysis of 60 cardiovascular secondary prevention trials (94,000 subjects) found that aspirin is associated with a significant 50% to 60% relative increase, respectively, in the risk of major extracranial bleeding.[67,68] For the outcomes of death and systemic embolism, pooled estimates of treatment effect from trials of antiplatelet monotherapy in patients with AF were imprecise, leaving uncertainty about the impact of antiplatelet monotherapy on these outcomes when compared with no treatment (Table 4).

2.1.3 VKAs vs Antiplatelet Monotherapy (Aspirin): The evidence summarized in sections 2.1.1 and 2.1.2 implies that adjusted-dose warfarin is far superior to aspirin for the prevention of stroke in patients with AF but is likely to be associated with a greater risk of bleeding complications. Direct evidence regarding this clinical question comes from 11 RCTs (total of 6,526 patients) comparing adjusted-dose VKA therapy to antiplatelet monotherapy (AFASAK 1, AFASAK 2, Birmingham Atrial Fibrillation Treatment of the Aged [BAFTA], EAFT, National Study for Prevention of Embolism in Atrial Fibrillation [NASPEAF], Primary Prevention of Arterial Thromboembolism in Nonrheumatic AF in Primary Care Trial [PATAF], Studio Italiano Fibrillazione Atriale [SIFA], SPAF II, SPAF III, Vemmos et al,[69] and Warfarin vs Aspirin for Stroke Prevention in Octogenarians With AF [WASPO]) (Table 5).[55,60,70-77] Antiplatelet therapy was typically with aspirin 75 to 325 mg/d, but in the SIFA and NASPEAF studies, it was with indobufen and triflusal, respectively.[72,74] In SPAF III and one of the two antiplatelet arms of AFASAK II, aspirin was given in combination with fixed minidose (ineffective) warfarin.[70,76]

These trials provide high-quality evidence that adjusted-dose VKA therapy reduces by one-half the risk of nonfatal stroke compared with antiplatelet monotherapy. These trials suggest that VKA therapy increases the risk of nonfatal major extracranial bleeding by about 50% compared with aspirin (pooled risk ratio, 1.42; 95% CI, 0.89-2.29), but the quality of evidence was rated down to moderate because of imprecision. Indirect evidence from RCTs of adjusted-dose VKA therapy vs aspirin in other populations suggest that VKA therapy is likely associated with a true twofold to 2.5-fold increase in major bleeding risk.[78,79] For the outcomes of death and systemic embolism, pooled estimates of treatment effect from trials of VKA therapy in patients with AF were imprecise, leaving uncertainty about the impact of VKA therapy on these outcomes compared with antiplatelet monotherapy.

2.1.4 VKAs vs Dual Antiplatelet Therapy With Aspirin and Clopidogrel: The Atrial Fibrillation Clopidogrel Trial With Irbesartan for Prevention of Vascular Events (ACTIVE) W trial assessed dual antiplatelet therapy with aspirin and clopidogrel as a potential alternative to VKA therapy (INR 2.0-3.0).[80] The trial was stopped early because of findings of superiority of VKA therapy (for their primary outcome of stroke, systemic embolism, MI, or vascular death) and did not find evidence of a difference in the risk of major bleeding (Table 6). Most patients (77%) were receiving VKA therapy before randomization, raising some concerns about generalizability of these findings to patients with AF who are being newly started on VKA therapy (ie, by enrolling mostly

Table 4—*[Section 2.1.2, 2.1.8] Should Antiplatelet (Aspirin) Monotherapy Rather Than No Therapy Be Used in Patients With AF?*[a,b]

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1 y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With No Therapy	With Antiplatelet	Relative Effect (95% CI)	With No Therapy	With Antiplatelet (95% CI)	Quality of Evidence
Death (critical outcome) mean follow-up 2.1 y											
6 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[c]	Undetected	231/2,152 (10.7)	215/2,246 (9.6)	RR 0.89 (0.75-1.05)	53 per 1,000[d]	6 fewer deaths per 1,000 (from 13 fewer to 3 more)	Moderate
Nonfatal stroke (critical outcome) mean follow-up 2.1 y; ischemic stroke and intracranial hemorrhage[e]											
8 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[f]	Undetected	197/2,274[g] (8.7)	174/2,498[h] (7.0)	RR 0.79 (0.65-0.96)	$CHADS_2$ 0 points		Moderate
									8 per 1,000[i]	2 fewer strokes per 1,000 (from 0 fewer to 3 fewer)	
									$CHADS_2$ 1 point		
									22 per 1,000	5 fewer strokes per 1,000 (from 1 fewer to 8 fewer)	
									$CHADS_2$ 2 points		
									45 per 1,000	9 fewer strokes per 1,000 (from 2 fewer to 16 fewer)	
									$CHADS_2$ 3-6 points		
									96 per 1,000	20 fewer strokes per 1,000 (from 4 fewer to 34 fewer)	
Nonfatal major extracranial bleeds (important outcome)[j]											
60 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	333/47,168 (0.7)	535/47,158 (1.1)	RR 1.60 (1.40-1.80)	5 per 1,000[k]	3 more bleeds per 1,000 (from 2 more to 4 more)	High

(Continued)

Table 4—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1 y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With No Therapy	With Antiplatelet	Relative Effect (95% CI)	With No Therapy	With Antiplatelet (95% CI)	Quality of Evidence
Systemic embolism (important outcome) mean follow-up 2.1 y											
5 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Precise[l]	Undetected	21/2,061 (1.0)	17/2,052 (0.8)	RR 0.80 (0.43-1.52)	4 per 1,000[m]	1 fewer systemic embolism per 1,000 (from 2 fewer to 2 more)	Moderate

ESPS = European Stroke Prevention Study; JAST = Japan Atrial Fibrillation Stroke Trial; LASAF = Low-dose Aspirin, Stroke, Atrial Fibrillation; SAFT = Swedish Atrial Fibrillation Trial; UK-TIA = United Kingdom Transient Ischaemic Attack Aspirin Trial. See Table 1-3 legends for expansion of other abbreviations.

[a]Pooled estimates of treatment effect in this evidence profile, with the exception of nonfatal major extracranial bleeding[b], are from a meta-analysis conducted for these guidelines, including data from RCTs of antiplatelet monotherapy (typically aspirin) vs no therapy (AFASAK I, SPAF I, EAFT, ESPS-2, UK-TIA, LASAF, JAST, SAFT).

[b]To obtain a best estimate of the risk of nonfatal major extracranial bleeding on antiplatelet monotherapy, we used data from the Antithrombotic Trialists' Collaboration[68] meta-analysis of aspirin for the secondary prevention of cardiovascular disease.

[c]The 95% CI does not exclude the possibility of no effect.

[d]From Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort).

[e]Intracranial hemorrhage includes intracerebral, subdural, and subarachnoid bleeds. For studies that did not report the number of strokes that were fatal and nonfatal (EAFT, ESPS-2, UK-TIA, JAST), we imputed values for the number of fatal and nonfatal strokes to estimate the pooled relative risk for nonfatal stroke across all eligible studies. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[f]Restricting meta-analysis exclusively to RCTs evaluating aspirin alone vs no antithrombotic therapy (ie, excluding SAFT, which used aspirin in combination with fixed minidose warfarin, and excluding the dipyridamole monotherapy arm of ESPS-2) results in an estimate of relative risk that includes no effect (relative risk, 0.81; 95% CI, 0.66-1.01).

[g]Of the 197 nonfatal strokes on no therapy, an estimated 193 (98%) were ischemic, and four (2%) were hemorrhagic.

[h]Of the 174 nonfatal strokes on antiplatelet therapy, an estimated 170 (98%) were ischemic, and four (2%) were hemorrhagic.

[i]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on no therapy by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[j]We used evidence from RCTs of aspirin for the secondary prevention of cardiovascular events to obtain a best estimate of the risk of major extracranial bleeding with antiplatelet monotherapy. Note that the study event rates and relative risk are for the outcome of all major extracranial bleeding (nonfatal and fatal events) because specific data regarding nonfatal events were not reported.

[k]Rates of nonfatal major extracranial bleeding on no therapy are extrapolated from rates observed in observational cohorts of patients receiving adjusted-dose VKA therapy (median, 1.3%/y) and dividing by the relative risk of major bleeding on adjusted-dose VKA therapy observed in the RCTs.

[l]The 95% CI does not exclude the possibility of appreciable harm or benefit with antiplatelet therapy.

[m]Based on rate of systemic embolism on aspirin of 0.3 per 100 patient-y reported in the IPD meta-analysis of warfarin vs aspirin by van Walraven et al[61] and a relative risk of 0.80 for aspirin vs no therapy for systemic embolism.

Table 5—*[Sections 2.1.3, 2.1.8, 2.1.9, 2.1.10] Should VKAs Rather Than Aspirin Be Used in Patients With AF?*[a]

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With ASA	With VKA	Relative Effect (95% CI)	With ASA	With VKA (95% CI)	Quality of Evidence
Death (critical outcome) mean follow-up 1.8 y											
10 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[b]	Undetected	351/3,020 (11.6)	327/2,835 (11.5)	RR 0.97 (0.85-1.12)	47 per 1,000[c]	1 fewer death per 1,000 (from 7 fewer to 6 more)	Moderate
Nonfatal stroke (critical outcome) mean follow-up 1.8 y; ischemic stroke and intracranial hemorrhage[d]											
11 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	209/3,356[e] (6.2)	94/3,170[f] (3.0)	RR 0.48 (0.33-0.70)	$CHADS_2$ 0 points		High
									6 per 1,000[g]	3 fewer strokes per 1,000 (from 2 fewer to 4 fewer)	
									$CHADS_2$ 1 point		
									17 per 1,000	9 fewer strokes per 1,000 (from 5 fewer to 11 fewer)	
									$CHADS_2$ 2 points		
									36 per 1,000	19 fewer strokes per 1,000 (from 11 fewer to 24 fewer)	
									$CHADS_2$ 3-6 points		
									76 per 1,000	40 fewer strokes per 1,000 (from 23 fewer to 51 fewer)	
Nonfatal major extracranial bleeds (important outcome) mean follow-up 1.8 y											
11 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[h]	Undetected	45/3,356 (1.3)	64/3,170 (2.0)	RR 1.42 (0.89-2.29)	8 per 1,000[i]	3 more bleeds per 1,000 (from 1 fewer to 10 more)	Moderate
Systemic embolism (important outcome) mean follow-up 1.8 y											
11 RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[j]	Undetected	20/3,356 (0.6)	13/3,170 (0.4)	RR 0.81 (0.40-1.64)	3 per 1,000[k]	1 fewer systemic embolism per 1,000 (from 2 fewer to 2 more)	Moderate

(Continued)

Table 5—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With ASA	With VKA	Relative Effect (95% CI)	With ASA	With VKA (95% CI)	Quality of Evidence
Burden of treatment (important outcome)											
N/A	No serious limitations	No serious limitations	No serious limitations	No serious limitations	N/A	N/A	N/A	VKA > aspirin	Warfarin: daily medication, lifestyle limitations, dietary restrictions, frequent blood testing and clinic visits Aspirin: daily medication only		High

ASA = acetylsalicyclic acid; BAFTA = Birmingham Atrial Fibrillation of the Aged; NASPEAF = National Study for Prevention of Embolism in Atrial Fibrillation; PATAF = Primary Prevention of Arterial Thromboembolism in Nonrheumatic AF in Primary Care Trial; SIFA = Studio Italiano Fibrillazione Artiale; WASPO = Warfarin vs Aspirin for Stroke Prevention in Octogenarians With AF. See Table 1-3 legends for expansion of other abbreviations.

[a]Pooled estimates of treatment effect in this evidence profile are from a new meta-analysis of data from RCTs which included a comparison of warfarin and antiplatelet monotherapy (typically aspirin). In SPAF III and AFASAK II, aspirin was combined with fixed (ineffective) minidose warfarin (AFASAK I, AFASAK II, Vemmos et al[69], BAFTA, EAFT, NASPEAF, PATAF, SIFA, SPAF II, SPAF III, WASPO).

[b]The 95% CI does not exclude the possibility of appreciable harm or benefit with VKA therapy.

[c]Based on data from Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort) and an estimated relative risk of mortality of 0.89 for aspirin vs no therapy (Table 4).

[d]Intracranial hemorrhage includes intracerebral, subdural, and subarachnoid bleeds. For SPAF II, which did not report the number of strokes that were fatal and nonfatal, we imputed values for the number of fatal and nonfatal strokes to estimate the pooled relative risk for nonfatal stroke across all eligible studies. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[e]Of the 209 nonfatal strokes on antiplatelet therapy, an estimated 202 (97%) were ischemic, and seven (3%) were hemorrhagic.

[f]Of the 94 nonfatal strokes on VKA therapy, an estimated 80 (85%) were ischemic, and 14 (15%) were hemorrhagic.

[g]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on no therapy by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[h]The 95% CI does not exclude appreciable harm with VKA therapy.

[i]Rates of nonfatal major extracranial bleeding on aspirin are extrapolated from rates observed in observational cohorts of patients receiving adjusted-dose VKA therapy (median, 1.3%/y), and relative risks of major extracranial bleeding associated with VKA therapy (RR, 2.56) and aspirin (RR, 1.60) compared with no therapy.

[j]The 95% CI does not exclude the possibility of appreciable benefit or harm with VKA therapy.

[k]Based on rate of systemic embolism on aspirin of 0.3 per 100 patient-y reported in the individual patient data meta-analysis of warfarin vs aspirin by van Walraven et al.[61]

Table 6—*[Sections 2.1.4, 2.1.9, 2.1.10] Should VKAs Rather Than Aspirin Plus Clopidogrel Be Used in Patients With AF?*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With ASA + Clopid	With VKA	Relative Effect (95% CI)	With ASA + Clopid	With VKA (95% CI)	Quality of Evidence
Death (critical outcome) median follow-up 1.3 y											
1 RCT	No serious limitations[a]	No serious inconsistency	No serious indirectness	Imprecise[b]	Undetected	159/3,335 (4.8)	158/3,371 (4.7)	RR 0.98 (0.79-1.22)	46 per 1,000[c]	1 fewer death per 1,000 (from 10 fewer to 10 more)	Moderate
Nonfatal stroke (critical outcome) median follow-up 1.3 y; ischemic stroke and intracranial hemorrhage[d]											
1 RCT	No serious limitations[a]	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	72/3,335[e] (2.2)	41/3,371[f] (1.2)	RR 0.56 (0.39-0.82)	$CHADS_2$ 0 points		High
									5 per 1,000[g]	2 fewer strokes per 1,000 (from 1 fewer to 3 fewer)	
									$CHADS_2$ 1 point		
									13 per 1,000	6 fewer strokes per 1,000 (from 2 fewer to 8 fewer)	
									$CHADS_2$ 2 points		
									26 per 1,000	11 fewer strokes per 1,000 (from 5 fewer to 16 fewer)	
									$CHADS_2$ 3-6 points		
									55 per 1,000	24 fewer strokes per 1,000 (from 10 fewer to 34 fewer)	
Nonfatal major extracranial bleeds (important outcome) median follow-up 1.3 y[h]											
1 RCT	No serious limitations[a]	No serious inconsistency	No serious indirectness (77% of patients were receiving VKA at study entry)	Imprecise[i]	Undetected	85/3,335 (2.5)	78/3,371 (2.3)	RR 0.91 (0.67-1.23)	12 per 1,000[j]	1 fewer bleeds per 1,000 (from 4 fewer to 3 more)	Moderate

(Continued)

Table 6—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With ASA + Clopid	With VKA	Relative Effect (95% CI)	With ASA + Clopid	With VKA (95% CI)	Quality of Evidence
Systemic embolism (important outcome) median follow-up 1.3 y											
1 RCT	No serious limitations[a]	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	18/3,335 (0.5)	4/3,371 (0.1)	RR 0.22 (0.07-0.65)	3 per 1,000[k]	2 fewer systemic emboli per 1,000 (from 1 fewer to 3 fewer)	High
Burden of treatment (important outcome)											
N/A	No serious limitations	No serious limitations	No serious limitations	No serious limitations	N/A	N/A	N/A	VKA > ASA + clopidogrel	Warfarin: daily medication, lifestyle limitations, dietary restrictions, frequent blood testing and clinic visits Aspirin + clopidogrel: daily medication only		High

ACTIVE = Atrial Fibrillation Clopidogrel Trial With Irbesartan for Prevention of Vascular Events. Clopid = clopidogrel. See Table 1-3 and 5 legends for expansion of other abbreviations.

[a]Results are based on a single RCT (ACTIVE W) that was stopped early for benefit.

[b]The 95% CI does not exclude appreciable harm or benefit with VKA therapy.

[c]Based on data from Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort), an estimated relative risk of 0.89 for mortality with aspirin vs no therapy (Table 4), and an estimated relative risk of 0.98 for mortality with aspirin + clopidogrel vs aspirin (Table 7).

[d]Intracranial hemorrhage includes intracerebral, subdural, and subarachnoid bleeds. The published report did not provide the number of strokes that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal strokes to estimate the relative risk for nonfatal stroke. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[e]Of the 72 nonfatal strokes on aspirin + clopidogrel therapy, an estimated 67 (93%) were ischemic, and five (7%) were hemorrhagic.

[f]Of the 41 nonfatal strokes on VKA therapy, an estimated 31 (76%) were ischemic, and 10 (24%) were hemorrhagic.

[g]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on aspirin + clopidogrel by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[h]The published report did not provide the number of major extracranial bleeds that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal major extracranial bleeds to estimate the relative risk for nonfatal major extracranial bleeds. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[i]The 95% CI does not exclude appreciable harm or benefit with VKA therapy.

[j]Rate of nonfatal major extracranial bleeding on aspirin + clopidogrel is extrapolated from rates observed in observational cohorts of patients receiving VKA therapy (median, 1.3%/y) and estimated relative risks of nonfatal major extracranial bleeding associated with VKA therapy compared with no therapy (RR, 2.56), aspirin compared with no therapy (RR, 1.60), and aspirin + clopidogrel compared with aspirin (RR, 1.50).

[k]Estimate is derived from rate of systemic embolism on aspirin (0.3 per 100 patient-y) reported in the IPD meta-analysis by van Walraven et al[61] and estimated relative risk of 0.97 (95% CI, 0.67-1.40) for systemic embolism with combination aspirin + clopidogrel therapy vs aspirin observed in ACTIVE A (Table 7).

prior users of VKA therapy, the study sample may be more representative of patients in whom VKA therapy is well tolerated). In prespecified subgroup analyses, the investigators did not find evidence of a difference in the effect of VKA therapy on primary outcome in patients who were and were not receiving VKA therapy at study entry. However, there was a significant difference (interaction $P = .03$) in the effect on major bleeding, depending on whether patients were prior users of VKA therapy. For patients who were not receiving VKA therapy at study entry, VKA therapy was associated with a nonsignificant trend toward a 69% relative increase in major bleeding compared with dual antiplatelet therapy, whereas in patients already receiving VKA therapy at study entry, VKA therapy was associated with a nonsignificant trend toward a 24% relative decrease in major bleeding compared with dual antiplatelet therapy.

2.1.5 Dual Antiplatelet Therapy With Aspirin and Clopidogrel vs Aspirin Alone: The ACTIVE A study compared combination aspirin and clopidogrel therapy with aspirin alone.[81] The trial enrolled 7,554 patients considered unsuitable for VKA therapy (approximately one-half because of a physician's judgment that VKA was inappropriate, one-fourth because of a specific risk of bleeding, and one-fourth because of the patient's preference not to take a VKA as the sole reason) and found that combination therapy is more effective in reducing the risk of nonfatal stroke in patients with AF but also increases the risk of nonfatal major extracranial bleeds compared with treatment with aspirin alone (Table 7).

2.1.6 New Oral Anticoagulants vs VKAs: Antithrombotic therapy for AF is evolving rapidly because of the development of new oral anticoagulants that directly target different parts of the coagulation pathway, have a more predictable anticoagulant effect, and do not require INR monitoring. Included in this new group of drugs are direct thrombin inhibitors (eg, dabigatran) and direct factor Xa inhibitors (apixaban, rivaroxaban, edoxaban). Results of large phase 3 clinical trials of these agents in patients with AF have been recently published or will be reported soon (Table 8). Although ximelagatran is no longer approved for use by regulatory agencies because of concerns about severe liver toxicity, the Stroke Prevention Using an Oral Thrombin Inhibitor in Patients With AF (SPORTIF) III and V trials were a proof of principle that a direct thrombin inhibitor can achieve similar protection against stroke compared with warfarin (ie, findings met the investigators' prespecified noninferiority criterion) with no evidence of increased bleeding risk.[82,83]

The Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) trial reported the results of a three-arm RCT of 18,113 patients with AF in which dabigatran 110 mg bid, and dabigatran 150 mg bid, were compared with open-label, adjusted-dose warfarin (target INR 2.0-3.0).[84] Based on best estimates of the proportion of strokes and major extracranial bleeds that were nonfatal (the published report did not present the number or proportion of fatal and nonfatal events), dabigatran at a dose of 150 mg bid is associated with a statistically significant one-third reduction in nonfatal stroke, with no evidence of a difference in the risk of nonfatal major extracranial bleeding compared with warfarin. Moreover, the data raised the possibility that dabigatran 150 mg bid may reduce all-cause mortality compared with warfarin (relative risk, 0.89; 95% CI, 0.79-1.01) (Table 9). In contrast, dabigatran at a dose of 110 mg bid was not associated with a significant difference in the risk of death, nonfatal stroke, nonfatal major extracranial bleeding, or systemic embolism (Table 10).

ROCKET-AF (Rivaroxaban Once Daily Oral direct Factor Xa inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation) was a double-blind, double-dummy RCT comparing rivaroxaban 20 mg once daily to adjusted-dose warfarin (INR 2.0-3.0) in 14,264 patients with AF at increased risk of stroke (mean $CHADS_2$ score of 3.5).[85] In the intention-to-treat analysis, rivaroxaban was noninferior to warfarin for the primary end point of stroke (ischemic or hemorrhagic) or systemic embolism but was not superior to warfarin (hazard ratio, 0.88; 95% CI, 0.74-1.03). The trial did not find evidence of a difference in major bleeding between rivaroxaban and warfarin (hazard ratio, 1.04; 95% CI, 0.90-1.20). Major GI bleeding was more common with rivaroxaban than with warfarin (3.2% and 2.2%, respectively, $P < .001$). Mortality was not significantly different between rivaroxaban and warfarin.

The Apixaban vs Acetylsalicylic Acid to Prevent Strokes (AVERROES) RCT compared apixaban 5 mg bid to aspirin in 5,599 patients with AF who were demonstrated or expected to be unsuitable candidates for adjusted-dose VKA therapy.[86] The trial was stopped early for benefit and was consistent with the trials of VKA therapy vs aspirin (section 2.1.3), reporting that apixaban reduces by one-half the occurrence of the primary outcome of stroke and systemic embolism compared with aspirin (hazard ratio, 0.45; 95% CI, 0.32-0.62). The trial failed to demonstrate or exclude a difference in the risk of major extracranial bleeding with apixaban compared with aspirin (hazard ratio, 1.23; 95% CI, 0.74-2.05), and the results were consistent with those of earlier trials of VKA therapy vs aspirin (section 2.1.3), which suggest an increase in the risk of major extracranial

Table 7—*[Sections 2.1.5, 2.1.8, 2.1.9, 2.1.10] Should Aspirin Plus Clopidogrel Rather Than Aspirin Be Used in Patients With AF?*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With ASA	With ASA/Clopid	Relative Effect (95% CI)	With ASA	With ASA/Clopid (95% CI)	Quality of Evidence
Death (critical outcome) median follow-up 3.6 y											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	841/3,782 (22.2)	825/3,772 (21.9)	RR 0.98 (0.90-1.07)	47 per 1,000[b]	1 fewer death per 1,000 (from 5 fewer to 3 more)	Moderate
Nonfatal stroke (critical outcome) median follow-up 3.6 y; ischemic stroke and intracranial hemorrhage[c]											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	315/3,782[d] (8.3)	226/3,772[e] (6.0)	RR 0.72 (0.61-0.85)	$CHADS_2$ 0 points		High
									6 per 1,000[f]	2 fewer strokes per 1,000 (from 1 fewer to 2 fewer)	
									$CHADS_2$ 1 point		
									17 per 1,000	5 fewer strokes per 1,000 (from 3 fewer to 7 fewer)	
									$CHADS_2$ 2 points		
									36 per 1,000	10 fewer strokes per 1,000 (from 5 fewer to 14 fewer)	
									$CHADS_2$ 3-6 points		
									76 per 1,000	21 fewer strokes per 1,000 (from 11 fewer to 30 fewer)	
Nonfatal major extracranial bleeds (important outcome) median follow-up 3.6 y[g]											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	112/3,782 (3.0)	167/3772 (4.4)	RR 1.50 (1.18-1.89)	8 per 1,000[h]	4 more bleeds per 1,000 (from 1 more to 7 more)	High

(Continued)

Table 7—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With ASA	With ASA/Clopid	Relative Effect (95% CI)	With ASA	With ASA/Clopid (95% CI)	Quality of Evidence
Systemic embolism (important outcome) median follow-up 3.6 y											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Precise[i]	Undetected	56/3,782 (1.5)	54/3,772 (1.4)	RR 0.97 (0.67-1.40)	3 per 1,000[j]	0 fewer systemic emboli per 1,000 (from 1 fewer to 1 more)	Moderate

See Table 1-3, 5, and 6 legends for expansion of abbreviations.

[a]The 95% CI does not exclude the possibility of no effect.

[b]Based on data from Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort) and an estimated relative risk of mortality of 0.89 for aspirin vs no therapy (Table 4).

[c]Hemorrhagic stroke not clearly defined in published RCT report (ie, unclear whether subdural and subarachnoid bleeds were included in the definition of hemorrhagic stroke).

[d]The published RCT report did not present the proportion of nonfatal strokes in the aspirin arm that were ischemic vs hemorrhagic; however, of all strokes (fatal and nonfatal), 5% were hemorrhagic, and 95% were ischemic or of uncertain type.

[e]The published RCT report did not present the proportion of nonfatal strokes in the combination aspirin + clopidogrel arm that were ischemic vs hemorrhagic; however, of all strokes (fatal and nonfatal), 10% were hemorrhagic, and 90% were ischemic or of uncertain type.

[f]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on aspirin by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[g]The published report did not provide the number of major extracranial bleeds that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal major extracranial bleeds to estimate the relative risk for nonfatal major extracranial bleeds. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[h]Rates of nonfatal major extracranial bleeding on aspirin are extrapolated from rates in observational cohorts of patients receiving adjusted-dose VKA therapy (median, 1.3%/y) and relative risks of major extracranial bleeding associated with VKA therapy (RR, 2.56) and aspirin (RR, 1.60) compared with no therapy.

[i]The 95% CI does not exclude the possibility of no effect.

[j]Based on rate of systemic embolism on aspirin of 0.3 per 100 patient-y reported in the IPD meta-analysis of warfarin vs aspirin by van Walraven et al.[61]

Table 8—*[Section 2.1.6] Phase 3 RCTs of New Oral Anticoagulants in Patients With AF*

Trial	Intervention	Comparator	Status
SPORTIF III SPORTIF V	Ximelagatran	Warfarin (INR, 2.0-3.0)	Published[82,83]
RE-LY	Dabigatran (150 or 110 mg bid)	Warfarin (INR, 2.0-3.0)	Published[84]
AVERROES	Apixaban (5 mg bid)	Aspirin (81-324 mg daily)	Published[86]
ROCKET-AF	Rivaroxaban (20 mg once daily)	Warfarin (INR, 2.0-3.0)	Published[85]
ARISTOTLE	Apixaban (5 mg bid)	Warfarin (INR, 2.0-3.0)	Published[87]
ENGAGE-AF TIMI 48	Edoxaban (high- and low-dose regimens)	Warfarin (INR, 2.0-3.0)	Currently recruiting

ARISTOTLE = Apixaban for the Prevention of Stroke in Subjects With Atrial Fibrillation; AVERROES = Apixaban versus Acetylsalicylic Acid to Prevent Strokes; ENGAGE-AF TIMI 48 = Effective Anticoagulation With Factor xA Next Generation in Atrial Fibrillation-Thrombolysis In Myocardial Infarction Study 48; INR = international normalized ratio; RE-LY = Randomized Evaluation of Long-Term Anticoagulation Therapy; ROCKET-AF = Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared With Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation; SPORTIF = Stroke Prevention Using an Oral Thrombin Inhibitor in Patients With AF.

bleeding with oral anticoagulation compared with aspirin.

ARISTOTLE (Apixaban for the Prevention of Stroke in Subjects With Atrial Fibrillation) was a double-blind, double-dummy RCT comparing apixaban 5 mg bid to warfarin (INR 2.0-3.0) in 18,201 patients with AF (mean $CHADS_2$ score of 2.1).[87] Apixaban reduced by 21% the risk of the primary outcome of stroke (ischemic or hemorrhagic) or systemic embolism (hazard ratio, 0.79; 95% CI, 0.66-0.95) and reduced by 31% the risk of major bleeding (hazard ratio, 0.69; 95% CI, 0.60-0.80) compared with warfarin. There was no evidence of a difference in major GI bleeding between apixaban and warfarin (hazard ratio, 0.89; 95% CI, 0.70-1.15). All-cause mortality was lower with apixaban compared with warfarin (hazard ratio, 0.89; 95% CI, 0.80-0.998). In all three recently completed trials of novel anticoagulants vs warfarin (RE-LY, ROCKET-AF, and ARISTOTLE), the rate of intracranial hemorrhage (including both hemorrhagic stroke and other intracranial bleeds) was lower in patients assigned to the novel anticoagulant than in patients assigned to warfarin.[84,85,87,88]

2.1.7 General Approach to Recommendations About New Oral Anticoagulants in This Article: Our guideline panel elected to make recommendations only for those drugs that have received regulatory approval for use in AF (ie, dabigatran).[89] Although based on the results of a single trial, there is evidence from RE-LY that dabigatran is no worse than VKA therapy with respect to nonfatal major extracranial bleeding and that it is similar or superior to warfarin with respect to nonfatal stroke, systemic embolism, and all-cause mortality in patients with nonvalvular AF (section 2.1.6). Therefore, for patients with nonrheumatic AF, wherever we recommend (or suggest) VKA therapy, we also recommend (or suggest) the use of dabigatran, and in these situations, our recommendations simply refer to oral anticoagulation.[88] We address the specific question of whether to use dabigatran over adjusted-dose VKA therapy in section 2.1.11.

The data from RE-LY do not directly address the use of dabigatran in patients with AF and mitral stenosis (patients with hemodynamically relevant valvular disease were excluded from this study), or in patients with AF in other special situations (sections 2.2 and 3.0). Therefore, we have not extrapolated the data from RE-LY to these clinical situations and have instead restricted those recommendations for oral anticoagulation to adjusted-dose VKA therapy. There is direct evidence from RE-LY regarding the use of dabigatran in patients with AF undergoing cardioversion, and these data are summarized in section 4.1.1.

2.1.8 Recommendations for Patients With AF at Low Risk of Stroke (eg, $CHADS_2$ Score of 0): Patients at sufficiently low risk of ischemic stroke may opt for no treatment rather than antithrombotic therapy with either aspirin or an oral anticoagulant. For instance, for every 1,000 patients at low risk of stroke with a $CHADS_2$ score of 0, VKA therapy compared with no treatment is anticipated to result in five fewer nonfatal strokes at the expense of eight more nonfatal major extracranial bleeds and the additional burden of adjusted-dose VKA treatment (Table 3). Although VKA therapy is expected to reduce all-cause mortality in patients with AF in general, it is likely that this mortality benefit does not extend to low-risk patients. The absolute reduction in fatal ischemic stroke with VKA therapy will be far fewer such patients, whereas their absolute increase in fatal intracranial hemorrhage will be similar to those with higher $CHADS_2$ scores.[51]

For patients with a $CHADS_2$ score of 0, treatment with aspirin for 1 year may result in the prevention of two nonfatal strokes per 1,000 patients (moderate-quality evidence due to imprecision) at the expense of three additional nonfatal major extracranial bleeds per 1,000 patients compared with no treatment (Table 4). If stroke rates are truly declining over time,

Table 9—*[Section 2.1.6, 2.1.11] Should Dabigatran 150 mg bid Rather Than VKAs be Used in Patients With AF?*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With VKA	With Dabi 150 mg	Relative Effect (95% CI)	With VKA	With Dabigatran 150 (95% CI)	Quality of Evidence
Death (critical outcome) median follow-up 2.0 y											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Serious imprecision[a]	Undetected	487/6,022 (8.1)	438/6,076 (7.2)	RR 0.89 (0.79- 1.01)	38 per 1,000[b]	4 fewer deaths per 1,000 (from 8 fewer to 0 more)	Moderate
Nonfatal stroke (critical outcome) median follow-up 2.0 y; ischemic stroke and intracranial hemorrhage[c]											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	149/6,022[d] (2.5)	101/6,076[e] (1.7)	RR 0.67 (0.52-0.86)	$CHADS_2$ 0 points		High
									3 per 1,000[f]	1 fewer strokes per 1,000 (from 0 fewer to 1 fewer)	
									$CHADS_2$ 1 point		
									8 per 1,000	3 fewer strokes per 1,000 (from 1 fewer to 4 fewer)	
									$CHADS_2$ 2 points		
									17 per 1,000	6 fewer strokes per 1,000 (from 2 fewer to 8 fewer)	
									$CHADS_2$ 3-6 points		
									36 per 1,000	12 fewer strokes per 1,000 (from 5 fewer to 17 fewer)	
Nonfatal major extracranial bleeds (important outcome) median follow-up 2.0 y[g]											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	264/6,022 (4.4)	286/6076 (4.7)	RR 1.07 (0.91-1.26)	13 per 1,000[h]	1 more bleed per 1,000 (from 1 fewer to 3 more)	Moderate
Systemic embolism (important outcome) median follow-up 2.0 y											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	14/6,022 (0.2)	12/6076 (0.2)	RR 0.85 (0.39-1.84)	2 per 1,000[i]	0 fewer systemic emboli per 1,000 (from 1 fewer to 2 more)	Moderate

(Continued)

Table 9—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With VKA	With Dabi 150 mg	Relative Effect (95% CI)	With VKA	With Dabigatran 150 (95% CI)	Quality of Evidence
Burden of treatment (important outcome)											
N/A	No serious limitations	No serious limitations	No serious limitations	No serious limitations	N/A	N/A	N/A	VKA >dabigatran	Warfarin: daily medication, lifestyle limitations, dietary restrictions, frequent blood testing and clinic visits Dabigatran: daily medication only		High

Dabi = dabigatran. See Table 1-3 legends for expansion of other abbreviations.

[a]The 95% CI does not exclude the possibility of no effect.

[b]Based on data from Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort) and an estimated relative risk of mortality of 0.72 for VKA vs no therapy (Table 3).

[c]Intracranial hemorrhage includes intracerebral, subdural, and subarachnoid bleeds. The published report did not provide the number of strokes that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal strokes to estimate the relative risk for nonfatal stroke. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[d]Of the 149 nonfatal strokes on VKA therapy, an estimated 106 (71%) were ischemic, and 43 (29%) were hemorrhagic.

[e]Of the 101 nonfatal strokes on dabigatran therapy, an estimated 83 (82%) were ischemic, and 18 (18%) were hemorrhagic.

[f]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on warfarin by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[g]The published report did not provide the number of major extracranial bleeds that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal major extracranial bleeds to estimate the relative risk for nonfatal major extracranial bleeds. Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[h]Rate of nonfatal major extracranial bleeding on warfarin is from observational cohorts of patients receiving adjusted-dose VKA therapy (median, 1.3%/y).

[i]Based on rate of systemic embolism on aspirin of 0.3 per 100 patient-y reported in the IPD meta-analysis of warfarin vs aspirin by van Walraven et al[61] and a relative risk for systemic embolism of 0.81 with VKA vs aspirin (Table 5).

Table 10—*[Section 2.1.6, 2.1.11] Should Dabigatran 110 mg bid Rather Than VKAs Be Used in Patients With AF?*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With VKA	With Dabi 110 mg	Relative Effect (95% CI)	With VKA	With Dabigatran 110 (95% CI)	Quality of Evidence
Death (critical outcome) median follow-up 2.0 y											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	487/6,022 (8.1)	446/6,015 (7.4)	RR 0.92 (0.81-1.04)	38 per 1,000[b]	3 fewer deaths per 1,000 (from 7 fewer to 2 more)	Moderate
Nonfatal stroke (critical outcome) median follow-up 2.0 y; ischemic stroke and intracranial hemorrhage[c]											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	149/6,022[d] (2.5)	132/6,015[e] (2.2)	RR 0.89 (0.70-1.12)	$CHADS_2$ 0 points		Moderate
									3 per 1,000[f]	0 fewer strokes per 1,000 (from 1 fewer to 0 more)	
									$CHADS_2$ 1 point		
									8 per 1,000	1 fewer strokes per 1,000 (from 2 fewer to 1 more)	
									$CHADS_2$ 2 points		
									17 per 1,000	2 fewer strokes per 1,000 (from 5 fewer to 2 more)	
									$CHADS_2$ 3-6 points		
									36 per 1,000	4 fewer strokes per 1,000 (from 11 fewer to 4 more)	
Nonfatal major extracranial bleeds (important outcome) median follow-up 2.0 y[g]											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	264/6,022 (4.4)	250/6,015 (4.2)	RR 0.95 (0.80-1.12)	13 per 1,000[h]	1 fewer bleeds per 1,000 (from 3 fewer to 2 more)	Moderate

(Continued)

Table 10—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With VKA	With Dabi 110 mg	Relative Effect (95% CI)	With VKA	With Dabigatran 110 (95% CI)	Quality of Evidence
Systemic embolism (important outcome) median follow-up 2.0 y											
1 RCT	No serious limitations	No serious inconsistency	No serious indirectness	Imprecise[a]	Undetected	14/6,022 (0.2)	11/6,015 (0.2)	RR 0.79 (0.36-1.73)	2 per 1,000[i]	0 fewer systemic emboli per 1,000 (from 1 fewer to 1 more)	Moderate
Burden of treatment (important outcome)											
N/A	No serious limitations	No serious limitations	No serious limitations	No serious limitations	N/A	N/A	N/A	VKA > dabigatran	Warfarin: daily medication, lifestyle limitations, dietary restrictions, frequent blood testing and clinic visits Dabigatran: daily medication only		High

See Table 1-3 and 9 legends for expansion of abbreviations.

[a]The 95% CI does not exclude the possibility of no effect.

[b]Based on data from Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort) and an estimated relative risk of mortality of 0.72 for VKA vs no therapy (Table 3).

[c]Intracranial hemorrhage includes intracerebral, subdural, and subarachnoid bleeds. The published report did not provide the number of strokes that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal strokes to estimate the relative risk for nonfatal stroke. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[d]Of the 149 nonfatal strokes on VKA therapy, an estimated 106 (71%) were ischemic, and 43 (29%) were hemorrhagic.

[e]Of the 132 nonfatal strokes on dabigatran therapy, an estimated 119 (90%) were ischemic, and 13 (10%) were hemorrhagic.

[f]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on warfarin by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[g]The published report did not provide the number of major extracranial bleeds that were fatal and nonfatal; therefore, we imputed values for the number of fatal and nonfatal major extracranial bleeds to estimate the relative risk for nonfatal major extracranial bleeds. Assumptions underlying these estimates are detailed in section 1.3 of this article.

[h]Rate of nonfatal major extracranial bleeding on warfarin is from observational cohorts of patients receiving adjusted-dose VKA therapy (median, 1.3%/y).

[i]Based on rate of systemic embolism on aspirin of 0.3 per 100 patient-y reported in the IPD meta-analysis of warfarin vs aspirin by van Walraven et al[61] and a relative risk for systemic embolism of 0.81 with VKA vs aspirin (Table 5).

then the already small benefits of antithrombotic therapy (number needed to treat for 1 year to prevent one nonfatal stroke of 500 for aspirin and 200 for VKA therapy) will be even smaller.

For patients who do choose antithrombotic therapy, the potential choices are aspirin, dual antiplatelet therapy with aspirin and clopidogrel, or oral anticoagulation. For every 1,000 patients with a $CHADS_2$ score of 0, treatment for 1 year with adjusted-dose VKA therapy or with combination aspirin and clopidogrel therapy compared with aspirin is anticipated to result in small reductions in nonfatal stroke and an increase in nonfatal major extracranial bleeding such that the net benefits of either VKA therapy or dual antiplatelet therapy with aspirin and clopidogrel would be small, particularly given the possibility of declining stroke rates over time (Tables 5, 7).

Recommendation

2.1.8. For patients with AF, including those with paroxysmal AF, who are at low risk of stroke (eg, $CHADS_2$ score = 0), we suggest no therapy rather than antithrombotic therapy (Grade 2B)**. For patients who do choose antithrombotic therapy, we suggest aspirin (75 mg to 325 mg once daily) rather than oral anticoagulation** (Grade 2B) **or combination therapy with aspirin and clopidogrel** (Grade 2B)**.**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with antithrombotic therapy are likely to choose antithrombotic therapy rather than no antithrombotic therapy. Other factors that may influence the choices above are a consideration of patient-specific bleeding risk and the presence of additional risk factors for stroke, including age 65 to 74 years and female sex, which have been more consistently validated, and vascular disease, which has been less well validated (see section 2.1.12). The presence of multiple non-$CHADS_2$ risk factors for stroke may favor oral anticoagulation therapy.

2.1.9 Recommendations for Patients With AF at Intermediate Risk of Stroke (eg, $CHADS_2$ Score of 1): For patients at intermediate risk of stroke with a $CHADS_2$ score of 1, compared with no therapy, 1 year of VKA therapy is expected to result in 15 fewer deaths and 15 fewer nonfatal strokes per 1,000 patients at the cost of eight more nonfatal major extracranial bleeds (Table 3). Regarding the choice between VKA therapy and aspirin, VKA therapy is anticipated to prevent nine nonfatal strokes for every 1,000 patients treated for 1 year compared with aspirin but will result in three additional bleeds and no reduction in all-cause mortality. However, because absolute rates of stroke may have fallen over the past 2 decades, we may be overestimating the absolute reduction in nonfatal stroke achieved with VKA therapy. Moreover, the true extent of bleeding risk with VKA therapy compared with aspirin therapy is unclear because the pooled estimate of the relative risk of bleeding from the relevant RCTs is imprecise (Table 5). The limited ability of $CHADS_2$ to accurately predict stroke risk (C statistic, 0.6-0.7), the considerable variability in patient values and preferences, and the burden and lifestyle limitations associated with adjusted-dose VKA therapy, introduce further uncertainty.

Compared with combination therapy with aspirin and clopidogrel, VKA therapy is expected to result in six fewer nonfatal strokes per 1,000 patients over a 1-year period, anywhere from four fewer to three more nonfatal major extracranial bleeds, and no reduction in all-cause mortality (Table 6). Uncertainty regarding the small net clinical benefit at a $CHADS_2$ score of 1 arises as a result of the limitations of the $CHADS_2$ score in estimating stroke risk and the possibility of declining absolute stroke rates over time. Uncertainty regarding the value of the small net benefit arises from the variability in patient values and preferences and the burden and lifestyle limitations associated with adjusted-dose VKA therapy.

For patients at intermediate risk of stroke with a $CHADS_2$ score of 1 who are unsuitable for or choose not to take an oral anticoagulant for reasons other than concerns about major bleeding (eg, difficulty maintaining a stable INR, lifestyle limitations of regular INR monitoring, dietary restrictions that are too burdensome, or costs of new anticoagulant drugs that are too high), combination therapy with aspirin and clopidogrel provides additional benefit of stroke reduction at the cost of additional bleeding (ACTIVE A trial) (Table 7). Patients opting for combination antiplatelet therapy rather than treatment with an oral anticoagulant should be informed that they are choosing an inferior treatment with regard to stroke prevention.

Recommendation

2.1.9. For patients with AF, including those with paroxysmal AF, who are at intermediate risk of stroke (eg, $CHADS_2$ score = 1), we recommend oral anticoagulation rather than no therapy (Grade 1B)**. We suggest oral anticoagulation rather than aspirin (75 mg to 325 mg once daily)** (Grade 2B) **or combination therapy with aspirin and clopidogrel** (Grade 2B)**. For patients who are unsuitable for or choose not to take an oral anticoagulant (for reasons other than concerns**

about major bleeding), we suggest combination therapy with aspirin and clopidogrel rather than aspirin (75 mg to 325 mg once daily) (Grade 2B)**.**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with anticoagulant therapy are likely to choose oral anticoagulation rather than antiplatelet therapy. Other factors that may influence the choice among antithrombotic therapies are a consideration of bleeding risk and the presence of additional risk factors for stroke, including age 65 to 74 years and female gender, which have been more consistently validated, and vascular disease, which has been less well validated (see section 2.1.12). The presence of multiple non-$CHADS_2$ risk factors for stroke may favor oral anticoagulation therapy.

2.1.10 Recommendations for Patients With AF at High Risk of Stroke (eg, $CHADS_2$ Score of ≥ 2, Which Includes Prior Ischemic Stroke or TIA): Patients at high risk of ischemic stroke, which includes patients with a history of ischemic stroke or TIA, can anticipate large benefits (ie, 15 fewer deaths and 30 fewer nonfatal strokes per 1,000 patients during 1 year of VKA therapy) with anticoagulation (Table 3).

For every 1,000 patients with a $CHADS_2$ score of 2 treated for 1 year with VKA therapy rather than aspirin, we anticipate 19 fewer nonfatal strokes at the expense of three more nonfatal major extracranial bleeds (Table 5). For every 1,000 such patients treated with VKA rather than combination therapy with aspirin and clopidogrel, we anticipate 11 fewer strokes and anywhere between four fewer to three more nonfatal major extracranial bleeds (Table 6). There is therefore a substantial net clinical benefit with oral anticoagulation.

For patients at high risk of ischemic stroke with a $CHADS_2$ score of ≥ 2 who are unsuitable for or who choose not to take an oral anticoagulant for reasons other than concerns about major bleeding (eg, difficulty maintaining a stable INR, lifestyle limitations of regular INR monitoring, dietary restrictions that are too burdensome, or costs of new anticoagulant drugs that are too high), aspirin and clopidogrel therapy will result in a substantial reduction in stroke compared with aspirin alone (ACTIVE A trial) (Table 7). Patients opting for combination antiplatelet therapy rather than treatment with an oral anticoagulant should be informed that they are choosing an inferior treatment with regard to stroke prevention.

Recommendation

2.1.10. For patients with AF, including those with paroxysmal AF, who are at high risk of stroke (eg, $CHADS_2$ score ≥ 2), we recommend oral anticoagulation rather than no therapy (Grade 1A)**, aspirin (75 mg to 325 mg once daily)** (Grade 1B)**, or combination therapy with aspirin and clopidogrel** (Grade 1B)**. For patients who are unsuitable for or choose not to take an oral anticoagulant (for reasons other than concerns about major bleeding), we recommend combination therapy with aspirin and clopidogrel rather than aspirin (75 mg to 325 mg once daily)** (Grade 1B)**.**

2.1.11 Recommendation Regarding Dabigatran vs Adjusted-Dose VKA Therapy: The RE-LY trial showed that dabigatran, at the higher dose of 150 mg bid, leads to reductions in nonfatal stroke, probable reductions in all-cause mortality, and no apparent increase in the risk of nonfatal major extracranial bleeding compared with VKA therapy (Table 9), whereas there was no evidence that dabigatran 110 mg bid leads to a significant reduction in relevant outcomes compared with VKA therapy (Table 10). In the United States, the Food and Drug Administration approved the use of dabigatran for the prevention of thromboembolism in patients with AF at a dose of 150 mg bid but not at a dose of 110 mg bid. However, the Food and Drug Administration did approve, based on pharmacokinetic considerations rather than direct evidence from RCTs in AF populations, a dose of 75 mg bid for patients with severe renal insufficiency (defined as a creatinine clearance 15-30 mL/min).[89]

For the question of whether to use dabigatran vs adjusted-dose VKA therapy, the evidence suggests net clinical benefit at the 150-mg dose. At the time of this writing, however, knowledge regarding the efficacy and safety of the new oral anticoagulants for patients with AF is still limited to one large randomized trial per agent. Uncommon but serious adverse effects may emerge with large-scale use of the drugs. Performance in usual clinical care may deteriorate because of less-restricted patient selection and suboptimal adherence to the unmonitored drug. For patients who do experience bleeding complications, clinicians need to be aware that there is no antidote to reverse the anticoagulant effects of dabigatran.[90] Given these concerns, it would be reasonable for VKA-experienced patients who are well controlled (ie, INR within therapeutic range a high proportion of the time) to continue on VKA therapy if they are satisfied with it and are tolerating it well rather than switching to dabigatran.

There is evidence from meta-analyses of RCTs[92] that home monitoring of VKA therapy reduces thromboembolic events by 42% compared with usual monitoring (see also Holbrook et al[91]), which is

similar to the 33% relative reduction in stroke achieved with dabigatran 150 mg bid compared with VKA therapy (Table 9).[84] Therefore, any advantages of dabigatran with respect to thromboembolism would likely not exist for motivated patients who are able to participate in home monitoring of their VKA therapy. Although home monitoring will reduce the burden of INR testing and VKA dose adjustment, the burdens of VKA therapy related to dietary restrictions and drug interactions will still exist, and there will be a cost for the home monitoring device and test strips. Therefore, depending on how patients value these burdens, some may choose dabigatran rather than home monitoring of VKA therapy.

Before prescribing dabigatran, clinicians need to judge whether the patient is similar enough to those enrolled in RE-LY that the clinical trial results are still likely to apply. In particular, the RE-LY study excluded patients with severe renal impairment (estimated creatinine clearance 30 mL/min or less).[84] The cost-effectiveness of the new anticoagulants compared with VKA therapy is another consideration. An economic analysis based on pricing of dabigatran in the United Kingdom (US $13 per day) estimated that dabigatran 150 mg bid would cost $45,372 more per quality-adjusted life year gained compared with warfarin for patients with AF aged 65 years with risk factors for stroke ($CHADS_2$ score of ≥ 1). The cost-effectiveness estimates in this model were sensitive to the pricing of dabigatran.[93]

Recommendation

2.1.11. For patients with AF, including those with paroxysmal AF, for recommendations in favor of oral anticoagulation (including 2.1.9, 2.1.10, and excluding 2.2, 3.1, 3.2, 3.3), we suggest dabigatran 150 mg twice daily rather than adjusted-dose VKA therapy (target INR range, 2.0-3.0) (Grade 2B)**.**

Remarks: Dabigatran is excreted primarily by the kidney. It has not been studied and is contraindicated in patients with severe renal impairment (estimated creatinine clearance of ≤ 30 mL/min). Clinicians should be aware that there is no antidote for dabigatran.

2.1.12 Tailoring These Recommendations to Individual Patients: As with all weak recommendations, treatment decisions should be individualized based on patient values and preferences, and in this case, an assessment of bleeding risk, and a consideration of additional risk factors for stroke.

Bleeding Risk Assessment—We have not made separate recommendations depending on patient bleeding risk because there are insufficient data to estimate reliably the absolute bleeding rates for patients in different categories of bleeding risk on different antithrombotic regimens. However, the following evidence regarding bleeding risk assessment may help to guide individualized treatment decisions for patients with AF.

One challenge in bleeding risk assessment for patients with AF is that many of the factors associated with an increased risk of major bleeding are also risk factors for ischemic stroke. For instance, older age, hypertension, congestive heart failure, and prior history of ischemic stroke—all of which are components of the $CHADS_2$ score—have been found in various studies of patients with AF to be independent predictors of bleeding while on VKA therapy.[94]

Several bleeding risk scores have been evaluated in cohorts of patients with AF (see Table 11).[53,95-98] However, these scores have not been extensively validated. Their ability to predict major bleeding is modest, with comparable C statistics across external validation studies (range, 0.61-0.66).[53,97-99] In scenarios where we make weak recommendations in favor of oral anticoagulation (eg, patients with a $CHADS_2$ score of 1), patients at high risk of major bleeding may decline oral anticoagulation.

Additional Risk Factors for Stroke in Patients With AF—In addition to consideration of bleeding risk and patient values and preferences, for patients with a $CHADS_2$ score of 0 or 1, clinicians may wish to consider additional stroke risk factors when individualizing decisions about antithrombotic therapy. For instance, there is high-quality evidence that stroke risk increases continuously with age (rather than as a dichotomous function of age <75 or ≥ 75 years) and moderate-quality evidence that female sex is an independent predictor of stroke risk in patients with AF.[18,20] Thus, patients may be inclined to choose the more aggressive treatment option (eg, antithrombotic therapy rather than no therapy for patients with a $CHADS_2$ score of 0 and oral anticoagulation rather than aspirin for patients with a $CHADS_2$ score of 1) when these additional risk factors for stroke are present. The presence of multiple non-$CHADS_2$ risk factors for stroke may favor oral anticoagulation therapy. There is less-consistent evidence supporting an independent association between vascular disease (ie, prior MI, complex aortic plaque seen on TEE, and peripheral arterial disease) and the risk of stroke in patients with AF (section 1.4.2).[21-23]

2.1.13 Percutaneous Closure of the Left Atrial Appendage: Percutaneous closure of the left atrial appendage (LAA) has recently been evaluated as a nondrug alternative for stroke prevention in patients with AF. The PROTECT-AF (WATCHMAN Left

Table 11—*[Section 2.1.12] Bleeding Risk Scores*

	Low	Moderate	High	Calculation of Bleeding Risk Score
Outpatient Bleeding Risk Index[95,96]	0	1-2	≥3	1 point for each of: Age ≥65 y GI bleed in past 2 wk Previous stroke Comorbidities (recent MI, Hct <30%, diabetes, or creatinine >1.5 mg/dL)
$HEMORR_2HAGES$[98]	0-1	2-3	≥4	1 point for each of: **H**epatic or renal disease **E**thanol abuse **M**alignancy **O**lder age (>75 yr) **R**educed platelet count or function **H**ypertension (uncontrolled) **A**nemia **G**enetic factors (*CYP2C9* polymorphisms) **E**xcessive fall risk **S**troke 2 points for: Rebleeding risk (ie, prior bleed)
Shireman et al[97]	≤1.07	>1.07 to <2.19	≥2.19	(0.49 × age >70) + (0.32 × female) + (0.58 × remote bleed) + (0.62 × recent bleed) + (0.71 × alcohol/drug abuse) + (0.27 × diabetes) + (0.86 × anemia) + (0.32 × antiplatelet drug use), with 1 point for presence of each and 0 if absent
HAS-BLED[53,99]	0	1-2	≥3	**H**ypertension (ie, uncontrolled BP) **A**bnormal renal/liver function (1 point each) **S**troke **B**leeding history or predisposition **L**abile INR **E**lderly (eg, age >65 y) **D**rugs (eg, concomitant antiplatelet/NSAID) or alcohol (1 point each) Maximum 9 points

Hct = hematocrit; NSAID = nonsteroidal antiinflammatory drug. See Table 1 and 8 legends for expansion of other abbreviations.

Atrial Appendage System for Embolic PROTECTion in Patients With Atrial Fibrillation) study randomized 707 patients with AF to percutaneous closure of the LAA using the WATCHMAN device (Atritech, Inc) or adjusted-dose warfarin to achieve a target INR of 2.0 to 3.0.[100] Percutaneous LAA closure was associated with a statistically nonsignificant reduction in the risk of their primary efficacy outcome of stroke (ischemic or hemorrhagic), cardiovascular or unexplained death, or systemic embolism compared with warfarin (absolute risk reduction, 1.9% per year; 95% CI, 3.2% per year less to 1.2% per year more, with percutaneous LAA closure compared with adjusted-dose warfarin). However, there was a significantly higher rate of adverse events (excessive bleeding, procedure-related complications) in the percutaneous LAA closure arm (absolute risk increase, 3.0% per year). In particular, serious pericardial effusion (ie, requiring percutaneous or surgical drainage) occurred in 4.8% of patients in the percutaneous LAA closure arm. Another occlusion device, PLAATO [percutaneous LAA transcatheter occlusion], has not been tested in a randomized trial but has been evaluated in prospective, multicenter cohort studies in patients ineligible for warfarin.[101,102] At this time, we make no formal recommendations regarding LAA closure devices, pending more definitive research in this field.

2.2 Patients With AF and Mitral Stenosis

Patients with AF in the setting of rheumatic mitral valve disease, particularly mitral stenosis, are at high risk of stroke. Most RCTs of adjusted-dose VKA therapy in AF excluded such patients. We believe that the results of randomized clinical trials in patients with nonrheumatic AF can be generalized to patients with mitral stenosis.

Recommendation

2.2. For patients with AF and mitral stenosis, we recommend adjusted-dose VKA therapy (target INR range, 2.0-3.0) rather than no therapy, aspirin (75 mg to 325 mg once daily), or combination therapy with aspirin and clopidogrel (all Grade 1B)**. For patients with AF and mitral stenosis who are unsuitable for or choose not to take adjusted-dose VKA therapy (for reasons other than concerns about major bleeding), we recommend combination therapy with aspirin**

and clopidogrel rather than aspirin (75 mg to 325 mg once daily) alone (Grade 1B)**.**

3.0 Antithrombotic Therapy for Patients With AF in Special Situations

3.1 Patients With AF and Stable Coronary Artery Disease

Approximately one-third of patients with AF also have coronary artery disease.[51] A recurring question is whether patients with AF for whom oral anticoagulation is indicated because of a high risk of stroke (eg, $CHADS_2$ score of ≥ 2) and who have concomitant stable coronary artery disease should also use aspirin to prevent coronary heart disease events. In this article, we define stable coronary artery disease as the presence (or absence) of angina but no revascularization procedure (percutaneous coronary intervention or coronary artery bypass graft surgery) or hospitalization for ACS (ie, unstable angina, non-ST-segment elevation MI, or ST-segment elevation MI) in the past year. The studies discussed next provide low-quality evidence that combination therapy with adjusted-dose VKA therapy plus aspirin is not associated with reductions in stroke or MI but that it does increase by 1.5 to 2 times the risk of major bleeding compared with adjusted-dose VKA therapy alone.

The FFAACS (Fluindione, Fibrillation Auriculaire, Aspirin et Contraste Spontane) trial is the only RCT that directly compared adjusted-dose VKA therapy and aspirin to adjusted-dose VKA therapy alone. In this study, patients in both arms received fluindione (INR 2.0-2.6), and patients in the combination therapy arm also received aspirin 100 mg/d. The trial was stopped early after enrollment of only 157 patients because of excessive hemorrhage in the group receiving fluindione plus aspirin.[103] No conclusion can be drawn regarding efficacy for prevention of stroke or MI because there were so few events during a short duration of follow-up.

There is direct evidence from a nonrandomized comparison of patients enrolled in the SPORTIF trials that combination therapy (with warfarin [INR 2.0-3.0] plus aspirin) is associated with a nearly twofold greater risk of major bleeding compared with warfarin alone, with no significant reduction in stroke or MI.[104] Patients receiving aspirin were different from those not receiving aspirin (eg, those receiving aspirin more often had diabetes, coronary artery disease, and previous stroke or TIA). Although the analyses were adjusted for baseline differences in patient characteristics, there remains a high risk for bias.

A similar nonrandomized comparison of patients enrolled in the RE-LY trial reported that rates (likely unadjusted) of major bleeding were roughly 2 times higher for patients receiving aspirin in conjunction with either warfarin (INR 2.0-3.0) or dabigatran.[105] Analyses of observational data from a large population-based registry of hospitalized patients with AF also found a nearly twofold increase in the risk of bleeding requiring hospitalization or causing death when patients with AF received combination therapy with warfarin and aspirin vs warfarin alone.[106]

A systematic review of RCTs in diverse (mostly non-AF) patient populations that compared aspirin plus VKA therapy with VKA therapy alone, in which VKAs were administered to achieve the same target INR or given at the same fixed dose in both arms, found that combination VKA plus aspirin therapy was associated with a lower risk of cardiovascular events compared with VKA therapy alone but that this benefit was restricted to RCTs enrolling patients with mechanical heart valves. No benefit was seen with combination therapy in studies of patients with AF or coronary artery disease, although estimates of treatment effect were very imprecise. Combination therapy with VKA and aspirin was associated with a greater risk of major bleeding (pooled OR, 1.43; 95% CI, 1.00-2.02) compared with VKA therapy alone.[107]

Recommendation

3.1. For patients with AF and stable coronary artery disease (eg, no acute coronary syndrome within the previous year) and who choose oral anticoagulation, we suggest adjusted-dose VKA therapy alone (target INR range, 2.0-3.0) rather than the combination of adjusted-dose VKA therapy and aspirin (Grade 2C)**.**

3.2 Patients With AF and Placement of an Intracoronary Stent (With or Without Recent ACS)

Patients benefit from dual antiplatelet therapy (eg, aspirin and clopidogrel) for a finite duration following placement of an intracoronary stent (4 weeks after placement of a bare-metal stent; 3 to 6 months after placement of a drug-eluting stent [typically 3 months for -olimus stents and 6 months for -taxel stents]) (see Vandvik et al).[108] The principal objective of dual antiplatelet therapy after placement of an intracoronary stent is the prevention of stent thrombosis. Before the adoption of dual antiplatelet therapy in clinical practice, stent thrombosis occurred in 6% to 24% of patients after bare-metal stent placement and was associated with a high case fatality rate of nearly 50%.[109-113] Concomitantly, a number of studies compared a new strategy, aspirin plus ticlopidine (a thienopyridine precursor to clopidogrel), to the previously most successful strategy of aspirin plus warfarin in patients undergoing stent placement. A Cochrane systematic review of four randomized

trials including 2,436 patients found that a 30- to 42-day course of ticlopidine plus aspirin vs warfarin plus aspirin reduced the 30- to 42-day risk of nonfatal MI (relative risk, 0.50; 95% CI, 0.30-0.83; number needed to treat, 55) and revascularization (relative risk, 0.29; 95% CI, 0.16-0.56; number needed to treat, 33) with a possible reduction in major bleeding (relative risk, 0.36; 95% CI, 0.14-1.02) (see Table 11 in Vandvik et al,[108] sections 3.1-3.5).[114]

Based on these data, we recommend aspirin and clopidogrel over warfarin plus aspirin for a finite period following stent placement. For patients with AF receiving oral anticoagulation who undergo placement of an intracoronary stent, the dilemma arises about whether patients should be continued on oral anticoagulation during the time they are recommended to be on dual antiplatelet therapy, given the lack of direct evidence from RCTs addressing this question.

Treatment decisions in this scenario must balance the effect of each drug combination on (1) the risk of stroke, systemic embolism, and mortality due to AF; (2) the risk of recurrent MI (including stent thrombosis); and (3) the risk of bleeding related to antithrombotic therapy. Because of the very high risk for bias in the available observational studies that compare cardiovascular event rates in patients receiving triple therapy vs dual antiplatelet therapy after stent placement,[115-120] we have instead used indirect evidence from relevant RCTs to inform our recommendations. We assumed that the relative impact of triple therapy on death, nonfatal MI, and nonfatal major extracranial bleeds compared with dual antiplatelet therapy would be similar to that seen in 10 RCTs of warfarin plus aspirin compared with aspirin in patients with ACS.[121] For nonfatal stroke and systemic embolism, we assumed that the relative impact of triple therapy vs dual therapy would be similar to that seen in the ACTIVE W trial, which compared warfarin vs aspirin plus clopidogrel in patients with AF (Table 6).[80] This assumption may be underestimating the true effect of triple therapy on stroke and systemic embolism compared with dual therapy.

We used the same annual baseline event rates (ie, on dual antiplatelet therapy) for death, nonfatal stroke, and systemic embolism as were used in the evidence profile comparing VKA to aspirin plus clopidogrel in the general AF population (Table 6). For the annual risk of nonfatal MI and nonfatal major extracranial bleeding on aspirin plus clopidogrel, we used the rates reported in the clopidogrel arm of the PLATO (Study of Platelet Inhibition and Patient Outcomes) trial, which compared ticagrelor to clopidogrel in patients with ACS receiving aspirin.[122]

The indirect data summarized in Table 12 suggest that triple therapy may be associated with net clinical benefit compared with dual antiplatelet therapy for patients at high risk of stroke (eg, $CHADS_2$ score of ≥ 2), whereas the net benefit at lower levels of stroke risk is uncertain. The duration of triple therapy should be kept as brief as possible given the associated increase in bleeding risk. Because the risk of stent thrombosis falls significantly after 30 days with bare-metal stents and 3 to 6 months after drug-eluting stents, triple therapy should be continued only during this high-risk period and only among patients at higher risk of stroke. After the initial period of triple therapy, patients may be given VKA therapy plus a single antiplatelet drug until 12 months have elapsed from the time of stent placement (particularly if stent placement was performed in the setting of a recent ACS [see section 3.3] or if a drug-eluting stent was used).

It should be noted that patients with AF who have received a drug-eluting stent and who are at increased risk of late stent thrombosis (eg, ACS at presentation, diabetes, long lesions, narrow diameter of target vessel)[123] may choose to continue triple therapy for a full 12 months after stent placement if they place a low value on avoiding bleeding. At 12 months after stent placement, antithrombotic therapy can be given according to our recommendations for AF and stable coronary artery disease (section 3.1).

Recommendation

3.2. For patients with AF at high risk of stroke (eg, $CHADS_2$ score of 2 or greater) during the first month after placement of a bare-metal stent or the first 3 to 6 months after placement of a drug-eluting stent, we suggest triple therapy (eg, VKA therapy, aspirin, and clopidogrel) rather than dual antiplatelet therapy (eg, aspirin and clopidogrel) (Grade 2C)**. After this initial period of triple therapy, we suggest a VKA (INR 2.0-3.0) plus a single antiplatelet drug rather than VKA alone** (Grade 2C)**. At 12 months after intracoronary stent placement, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

For patients with AF at low to intermediate risk of stroke (eg, $CHADS_2$ score of 0 or 1) during the first 12 months after placement of an intracoronary stent (bare metal or drug eluting), we suggest dual antiplatelet therapy rather than triple therapy (Grade 2C)**. At 12 months after intracoronary stent placement, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding

Table 12—*[Sections 3.2, 3.3] Should Triple Therapy Rather Than Dual Antiplatelet Therapy Be Used in Patients With AF After Intracoronary Stent Placement?*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With Dual Therapy	With Triple Therapy	Relative Effect (95% CI)	With Dual Antiplatelet Therapy	With Triple Therapy (95% CI)	Quality of Evidence
Death (critical outcome)											
10 RCTs	No serious limitations	No serious inconsistency	Very serious indirectness[a]	Imprecise[b]	Undetected			RR 1.00 (0.82-1.22)	46 per 1,000[c]	0 more deaths per 1,000 (from 8 fewer to 10 more)	Low
Nonfatal stroke (critical outcome) median follow-up 1.3 y; ischemic stroke and intracranial hemorrhage											
1 RCT	No serious limitations	No serious inconsistency	Very serious indirectness[d]	No serious imprecision	Undetected	72/3,335 (2.2)	41/3,371 (1.2)	RR 0.56 (0.39-0.82)	$CHADS_2$ 0 points		Low
									5 per 1,000[e]	2 fewer strokes per 1,000 (from 1 fewer to 2 fewer)	
									$CHADS_2$ 1 point		
									13 per 1,000	6 fewer strokes per 1,000 (from 2 fewer to 8 fewer)	
									$CHADS_2$ 2 points		
									26 per 1,000	11 fewer strokes per 1,000 (from 5 fewer to 16 fewer)	
									$CHADS_2$ 3-6 points		
									55 per 1,000	24 fewer strokes per 1,000 (from 10 fewer to 34 fewer)	
Nonfatal MI (important outcome)											
10 RCTs	No serious limitations	No serious inconsistency	Very serious indirectness[a]	No serious imprecision	Undetected			RR 0.69 (0.54-0.88)	69 per 1,000[f]	21 fewer MI per 1,000 (from 8 fewer to 32 fewer)	Low
Nonfatal major extracranial bleeds (important outcome)											
10 RCTs	No serious limitations	No serious inconsistency	Very serious indirectness[a]	No serious imprecision	Undetected			RR 2.37 (1.62-3.47)	19 per 1,000[g]	26 more bleeds per 1,000 (from 12 more to 47 more)	Moderate
Systemic embolism (important outcome) median follow-up 1.3 y											
1 RCT	No serious limitations	No serious inconsistency	Very serious indirectness[d]	No serious imprecision	Undetected	18/3,335 (0.5)	4/3,371 (0.1)	RR 0.22 (0.07-0.65)	3 per 1,000[h]	2 fewer systemic emboli per 1,000 (from 1 fewer to 3 fewer)	Low

(Continued)

Table 12—*Continued*

Quality Assessment								Summary of Findings			
						Study Event Rates (%)			Estimation of Absolute Effects 1-y Time Frame		
No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	With Dual Therapy	With Triple Therapy	Relative Effect (95% CI)	With Dual Antiplatelet Therapy	With Triple Therapy (95% CI)	Quality of Evidence
Burden of treatment (important outcome)											
N/A	No serious limitations	No serious limitations	No serious limitations	No serious limitations	N/A	N/A	N/A	Triple therapy > dual therapy	Triple therapy: daily medication, lifestyle limitations, dietary restrictions, frequent blood testing and clinic visits Dual antiplatelet therapy: daily medication only		High

PLATO = Study of Platelet Inhibition and Patient Outcomes; TIMI = Thrombolysis In Myocardial Infarction. See Table 1-3 legends for expansion of other abbreviations.

[a]We used the relative risk for all-cause mortality, nonfatal MI, and nonfatal major extracranial bleeding associated with warfarin + aspirin therapy vs aspirin alone in 10 RCTs of patients with acute coronary syndrome (Rothberg et al[121]) to estimate the effect of triple therapy vs dual antiplatelet therapy on all-cause mortality in patients with AF and acute coronary syndrome.

[b]The 95% CI does not exclude important harm or benefit with triple therapy.

[c]We used the same baseline estimates of mortality as for the general AF population receiving aspirin + clopidogrel. This was based on data from Go et al,[44] which reported a mortality rate of 5.33 per 100 person-y in untreated (no warfarin) patients from an observational study (ATRIA cohort), an estimated relative risk of 0.89 for mortality with aspirin vs no therapy (Table 4), and an estimated relative risk of 0.98 for mortality with aspirin + clopidogrel vs aspirin (Table 7).

[d]We used data from an RCT of warfarin vs aspirin + clopidogrel in patients with AF (ACTIVE W) to estimate the relative effect of triple therapy vs dual antiplatelet therapy on nonfatal stroke and systemic embolism in patients with AF and acute coronary syndrome. This assumption may be underestimating the potential benefit of triple therapy vs dual antiplatelet therapy in reducing nonfatal stroke.

[e]We estimated the annual rates of nonfatal stroke (ischemic and hemorrhagic) on aspirin + clopidogrel by extrapolating from observed annual rates of ischemic stroke (fatal + nonfatal), stratified by $CHADS_2$ score, in the aspirin arms of six historical RCTs in patients with AF (Gage et al[37]). Assumptions underlying these estimates are detailed in section 1.1.4 of this article.

[f]From clopidogrel (plus aspirin) arm of the PLATO trial of ticagrelor vs clopidogrel in patients with acute coronary syndrome.

[g]Based on the absolute rate of non-coronary artery bypass graft-related major bleeding (TIMI definition) reported in the clopidogrel (plus aspirin) arm of the PLATO trial minus the rate of fatal bleeding reported in this arm of the PLATO trial.

[h]Estimate is derived from rate of systemic embolism on aspirin (0.3 per 100 patient-y) reported in the individual patient data meta-analysis by van Walraven et al[61] and estimated relative risk of 0.97 (95% CI, 0.67-1.40) for systemic embolism with combination aspirin + clopidogrel therapy vs aspirin observed in ACTIVE A (Table 7).

bleeding and the burden associated with anticoagulant therapy are likely to choose triple therapy rather than dual antiplatelet therapy. Other factors that may influence this choice are a consideration of bleeding risk and the presence of additional non-$CHADS_2$ risk factors for stroke (see section 2.1.12).

3.3 Patients With AF and ACS Who Do Not Undergo Intracoronary Stent Placement

Dual antiplatelet therapy (eg, with aspirin and clopidogrel) rather than aspirin alone is recommended during the first 12 months after an ACS, regardless of whether patients also receive an intracoronary stent (Vandvik et al[108]). Many patients with AF (eg, $CHADS_2$ score of ≥ 1) will choose VKA therapy (eg, warfarin) to prevent future stroke (section 2.1). Therefore, treatment options in patients with AF and recent ACS may include warfarin plus dual antiplatelet therapy (ie, triple therapy), dual antiplatelet therapy, or warfarin plus single antiplatelet therapy. In this section, we specifically address patients with AF and ACS who do not undergo placement of an intracoronary stent.

The indirect data summarized in Table 12 suggest that triple therapy may not be associated with net clinical benefit compared with dual antiplatelet therapy unless patients are at substantially high risk of stroke (eg, $CHADS_2$ score of ≥ 2), whereas the net benefit at lower levels of stroke risk is uncertain. Rather than triple therapy, a third therapeutic option in this clinical scenario (where stent thrombosis is not a concern) is to use a VKA plus single antiplatelet therapy. Unfortunately, there are no RCTs comparing warfarin plus single antiplatelet therapy to dual antiplatelet therapy. As described in Author et al (section 2.4) in this guideline, warfarin plus aspirin is associated with a significant reduction in risk of subsequent MI (relative risk, 0.69; 95% CI, 0.54-0.88) and stroke (relative risk, 0.56; 95% CI, 0.39-0.82) compared with aspirin alone in patients post-ACS.[121] The point estimates for these reductions in MI and stroke are greater than those seen with clopidogrel plus aspirin (relative risk, 0.77; 95% CI, 0.67-0.89) vs aspirin alone (relative risk, 0.86; 95% CI, 0.63-1.18) in the Clopidogrel in Unstable Angina to Prevent Recurrent Events (CURE) trial.[124] These data suggest that warfarin plus aspirin is at least as effective, and potentially more effective, than clopidogrel plus aspirin for the prevention of cardiovascular events after ACS. Warfarin plus aspirin was not given a recommendation in Author et al either as an alternative or in preference to dual antiplatelet therapy for patients post-ACS because of pragmatic issues (physician reluctance, burden of use, etc). However, these issues are not relevant for patients already receiving warfarin for AF.

Finally, there are no studies comparing warfarin plus aspirin to triple therapy (eg, warfarin, aspirin, and clopidogrel) in patients with AF and recent ACS. Use of triple therapy in this situation would be appropriate only if the risk reduction in MI and stroke achieved by adding the second antiplatelet agent is greater than the increase in bleeding risk. This seems unlikely in patients who are not undergoing stent placement and, thus, where there is no concern regarding stent thrombosis. Given the lack of direct evidence and mindful of the principle to first do no harm, we do not advocate the use of triple therapy in patients with AF who experience ACS but do not receive an intracoronary stent. However, patients placing a high value on MI and stroke reduction and a low value on avoiding bleeding may opt for an initial period of triple therapy (eg, 3-6 months) followed by warfarin plus aspirin.

Recommendation

3.3. For patients with AF at intermediate to high risk of stroke (eg, $CHADS_2$ score of 1 or greater) who experience an acute coronary syndrome and do not undergo intracoronary stent placement, we suggest, for the first 12 months, adjusted-dose VKA therapy (INR 2.0-3.0) plus single antiplatelet therapy rather than dual antiplatelet therapy (eg, aspirin and clopidogrel) or triple therapy (eg, warfarin, aspirin, and clopidogrel) (Grade 2C)**. After the first 12 months, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

For patients with AF at low risk of stroke (eg, $CHADS_2$ score of 0), we suggest dual antiplatelet therapy (eg, aspirin and clopidogrel) rather than adjusted-dose VKA therapy (INR 2.0-3.0) plus single antiplatelet therapy or triple therapy (eg, warfarin, aspirin, and clopidogrel) (Grade 2C)**. After the first 12 months, antithrombotic therapy is suggested as for patients with AF and stable coronary artery disease (see section 3.1).**

Remarks: Patients who place an exceptionally high value on stroke reduction and a low value on avoiding bleeding and the burden associated with anticoagulant therapy are likely to choose adjusted-dose VKA therapy plus single antiplatelet therapy rather than dual antiplatelet therapy. Other factors that may influence this choice are a consideration of bleeding risk and the presence of additional non-$CHADS_2$ risk factors for stroke (see section 2.1.12).

3.4 Patients With AF Managed by a Rhythm Control Strategy

Some patients with AF will be managed with antiarrhythmic drugs (AADs) to achieve and maintain normal sinus rhythm. Increasingly, patients with AF are also receiving catheter radiofrequency ablation procedures (pulmonary vein isolation) to maintain normal sinus rhythm. In the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) trial, in which many patients in the rhythm control arm did not receive anticoagulation, patients in a rhythm control strategy had a similar risk of stroke compared with patients in a rate control strategy (see also section 4.1.1).[125]

There are several RCTs of catheter ablation vs AAD therapy in patients with AF, typically those who have failed first-line therapy with AADs.[126-132] All trials found a significant reduction in AF recurrence at ~1 year of follow-up. However, AF recurrence in the catheter ablation arms ranged from 11% to 44% at ~1 year. The studies rarely reported on stroke outcomes, and all were underpowered to address this question. Given the results of the AFFIRM trial, the lack of longer-term follow-up data from catheter ablation RCTs regarding AF recurrence rates, and poor reporting of stroke outcomes, it would be prudent to base decisions about long-term antithrombotic therapy on a patient's underlying risk for stroke as recommended in section 2.1, and not on their underlying rhythm.

Recommendation

3.4. For patients with AF being managed with a rhythm control strategy (pharmacologic or catheter ablation), we suggest that antithrombotic therapy decisions follow the general risk-based recommendations for patients with AF in section 2.1, regardless of the apparent persistence of normal sinus rhythm (Grade 2C)**.**

3.5 Patients With Atrial Flutter

Many patients with persistent atrial flutter have periods of atrial flutter alternating with periods of AF.[133,134] The prevalence of thrombus in the body of the atria and atrial appendage on TEE in patients with atrial flutter ranges from 1% to 21%.[135-138] There are few data from longitudinal studies assessing the risk of thromboembolism with well-documented sustained atrial flutter. A study describing a series of 191 consecutive, unselected patients hospitalized for treatment of atrial flutter reported thromboembolism in 7% of patients during a mean follow-up of 26 months.[134] The role of anticoagulant therapy for patients with atrial flutter has not been evaluated in clinical trials, but because these patients often have concomitant AF or are at increased risk of developing AF, it is reasonable to base decisions regarding antithrombotic therapy on the risk stratification schemes used for AF (section 1.1).

Recommendation

3.5. For patients with atrial flutter, we suggest that antithrombotic therapy decisions follow the same risk-based recommendations as for AF.

4.0 Antithrombotic Therapy for Patients With AF Undergoing Cardioversion

To minimize the risk of stroke and systemic embolism associated with cardioversion, therapeutic anticoagulation (eg, with adjusted-dose oral VKAs; INR 2.0-3.0) conventionally is recommended for a minimum of 3 weeks before, during, and for a minimum of 4 weeks after the procedure. For some patients with AF of documented short duration (eg, ≤48 h), a common practice is to cardiovert without prolonged precardioversion anticoagulation. For patients with AF duration of >48 h or unknown duration, a TEE-guided approach is an alternative strategy that can simplify anticoagulation management before cardioversion. In this section, we summarize the evidence and give recommendations for the use of antithrombotic therapy in patients undergoing electrical or pharmacologic cardioversion for AF (or atrial flutter).

4.1 Patients Undergoing Elective Cardioversion of AF

4.1.1 Cardioversion of AF of More Than 48 h or Unknown Duration: Evidence favoring the efficacy of pericardioversion anticoagulation is based on observational studies in mostly patients undergoing electrical rather than pharmacologic cardioversion. There is moderate-quality evidence from a systematic review of 18 observational studies suggesting that the risk of stroke or thromboembolism is substantially lower in patients receiving pericardioversion anticoagulation than in those who receive no anticoagulation (0.3% vs 2.0%), translating to a relative risk of 0.16 (95% CI, 0.05-0.48) in favor of anticoagulation.[139] No data regarding major hemorrhagic events were reported in this systematic review.

The conventional duration of a minimum of 3 weeks therapeutic anticoagulation before cardioversion and a minimum 4 weeks afterward is based on indirect pathophysiologic data and evidence from observational studies and remains arbitrary. Observational data showing that thromboembolism was significantly more common at an INR of 1.5 to 2.4 before

cardioversion than an INR of ≥ 2.5 (0.93% vs 0%, $P = .012$) suggests the importance of maintaining a therapeutic INR in the pericardioversion period.[140] After cardioversion, results of observational studies suggest that the highest risk of stroke and thromboembolism is in the first 72 h after cardioversion and that the majority of thromboembolic complications will occur within 10 days of cardioversion.[141] However, TEE studies have demonstrated that despite restoration of sinus rhythm on the ECG, atrial mechanical dysfunction may persist for several weeks postcardioversion.[142]

A nonrandomized comparison of 1,270 patients enrolled in the RE-LY trial who underwent 1,983 cardioversions suggests that there may be no excess harm with dabigatran compared with warfarin when used for pericardioversion anticoagulation. Most cardioversions in this study (~80%) were performed after the protocol-assigned study anticoagulant was given for a minimum of 3 weeks before the procedure, and rates of stroke and systemic embolism at 30 days after cardioversion were low when oral anticoagulation with either warfarin or dabigatran was given before cardioversion (0.8%, 0.3%, and 0.6% with dabigatran 110 mg bid, dabigatran 150 mg bid, and warfarin, respectively).[143]

A TEE-guided approach with abbreviated anticoagulation before cardioversion is an alternative to the conventional approach of using a minimum of 3 weeks therapeutic precardioversion anticoagulation.[144] Under a TEE-guided strategy, patients receive anticoagulation and once therapeutic, undergo screening TEE. If thrombus is seen in either atrial appendage or atrium at the time of TEE, cardioversion is postponed, given the presumed high risk of thromboembolism. If no thrombus is seen, the patient proceeds immediately to cardioversion. A TEE-guided strategy requires an experienced echocardiographer because accurate visualization of thrombus may be operator dependent.

The Assessment of Cardioversion Using Transesophageal Echocardiography (ACUTE) RCT compared a TEE-guided strategy of abbreviated therapeutic anticoagulation with IV unfractionated heparin (started 24 h before cardioversion) or warfarin (INR 2.0-3.0) (started 5 days before cardioversion) to a strategy of therapeutic anticoagulation for at least 3 weeks before cardioversion.[145] The evidence is of low quality given the wide 95% CIs around the point estimates of effect on patient-important outcomes (Table 13). Thus, the results do not exclude with confidence the possibility of important benefit or harm with a TEE-guided approach compared with a conventional approach of 3 weeks anticoagulation precardioversion.

Acknowledging these uncertainties, a TEE-guided approach may be best suited for patients who are very symptomatic while in AF because cardioversion can be done sooner if the TEE is negative for thrombus. It may also suit patients who would prefer to avoid prolonged oral anticoagulation before cardioversion and those at increased risk for bleeding. However, the ability to avoid anticoagulation with a TEE-guided strategy is most relevant for patients without stroke risk factors and at low risk of recurrent AF in whom long-term anticoagulation beyond 4 weeks after cardioversion would not be required.

For patients undergoing a TEE-guided approach, low-molecular-weight heparin at full VTE treatment doses or IV unfractionated heparin (to maintain an activated partial thromboplastin time prolongation

Table 13—*[Section 4.1] Abbreviated Anticoagulation With TEE-Guided Cardioversion vs Conventional Anticoagulation for at Least 3 Weeks Before Cardioversion*

	Anticipated Absolute Effects (Time Frame Is 8 wk)				
Outcomes	Risk With Conventional Anticoagulation	Risk Difference With TEE + Abbreviated Anticoagulation[a] (95% CI)	Relative Effect (95% CI)	No. of Participants (Studies), Follow-up	Quality of The Evidence (GRADE)
Death	10 per 1,000[b]	14 more deaths per 1,000 (0 more to 52 more)	RR 2.44 (0.95-6.24)	1,222 (1 RCT), 8 wk	Low[c]
Nonfatal strokes	3 per 1,000[b]	5 more strokes per 1,000 (2 fewer to 38 more)	RR 2.44 (0.47-12.50)	1,222 (1 RCT), 8 wk	Low[c]
Nonfatal major extracranial bleeds[d]	15 per 1,000[b]	7 fewer bleeds per 1,000 (12 fewer to 9 more)	RR 0.54 (0.18-1.61)	1,222 (1 RCT), 8 wk	Low[c]

ACUTE = Assessment of Cardioversion Using Transesophageal Echocardiography; GRADE = Grades of Recommendations, Assessment, Development, and Evaluation; TEE = transesophageal echocardiography. See Tables 1 and 8 for expansion of other abbreviations.

[a]Abbreviated anticoagulation refers to either IV unfractionated heparin (started 24 h before cardioversion) or adjusted-dose warfarin (INR 2.0-3.0) (started 5 d before cardioversion).

[b]Assumed risk is the observed event rate from the control arm (ie, conventional anticoagulation) of the RCT by Klein et al[145] (ACUTE trial).

[c]No statistically significant effect, and CIs are very wide. Trial was not sufficiently large to demonstrate comparable safety between conventional and TEE-based strategies.

[d]Most of these were GI bleeds.

that corresponds to plasma heparin levels of 0.3-0.7 International Units/mL antifactor Xa activity) should be started at the time of TEE and cardioversion performed within 24 h of the TEE if no thrombus is seen. A few observational studies and one RCT have suggested that low-molecular-weight heparin has similar efficacy compared with heparin or warfarin for immediate anticoagulation before TEE.[146-150] Outpatients undergoing a TEE-guided approach may be started on a VKA (INR 2.5; range, 2.0-3.0) and the TEE and subsequent cardioversion scheduled for 5 days later (if the INR is in therapeutic range at that time). The new oral anticoagulants may also be suitable for outpatient treatment before TEE-guided cardioversion given their ease of use (eg, dabigatran achieves steady-state concentrations in 2-3 days after bid administration),[151] but they have not yet been studied for this purpose.

There is no direct evidence to guide decisions about the long-term management of anticoagulation in patients who appear to be in sinus rhythm 4 weeks after cardioversion. Several observational studies indicate that approximately one-half of patients will have recurrence of AF at 1 year after cardioversion.[152-156] The AFFIRM study, in which many patients stopped anticoagulation after initial (apparently) successful restoration of sinus rhythm, found similar rates of thromboembolism with a rhythm control strategy compared with a rate control strategy.[125] Finally, patients with PAF often are asymptomatic during episodes of AF recurrence, with one series suggesting that only one in every 12 paroxysms are symptomatic.[157] These observations suggest that decisions about long-term antithrombotic therapy should be primarily based on a patient's risk for stroke (see section 2.1) rather than on the prevailing rhythm at 4 weeks postcardioversion.

Recommendation

4.1.1. For patients with AF of greater than 48 h or unknown duration undergoing elective electrical or pharmacologic cardioversion, we recommend therapeutic anticoagulation (adjusted-dose VKA therapy, target INR range 2.0-3.0, low-molecular-weight heparin at full venous thromboembolism treatment doses, or dabigatran) for at least 3 weeks before cardioversion or a TEE-guided approach with abbreviated anticoagulation before cardioversion, rather than no anticoagulation (Grade 1B)**. We recommend therapeutic anticoagulation for at least 4 weeks after successful cardioversion to sinus rhythm, rather than no anticoagulation, regardless of the baseline risk of stroke** (Grade 1B)**. Decisions about anticoagulation beyond 4 weeks should be made in accordance with our risk-based recommendations for long-term antithrombotic therapy in section 2.1.**

4.1.2 Cardioversion of AF of 48 h Duration or Less: There is uncertainty over the precise duration of AF necessary for thrombus to develop and, hence, the threshold of AF duration below which precardioversion anticoagulation can be safely avoided. For AF of short duration (eg, ≤48 h), a common practice is to cardiovert without a TEE or prolonged precardioversion anticoagulation. However, observational studies have found left atrial thrombus on TEE in as many as 14% of patients with acute AF of short duration.[158,159] Moreover, because many individuals develop AF asymptomatically, it is often difficult to accurately determine a patient's duration of AF, making the 48-h rule difficult to apply.[160] No RCTs have compared different anticoagulation strategies in patients with AF of documented duration of ≤48 h. Observational data suggest that the risk of stroke or thromboembolism in these patients is similar to those who have received conventional anticoagulation for a minimum of 3 weeks before cardioversion.[140,161]

Recommendation

4.1.2. For patients with AF of documented duration of 48 h or less undergoing elective cardioversion (electrical or pharmacologic), we suggest starting anticoagulation at presentation (low-molecular-weight heparin or unfractionated heparin at full venous thromboembolism treatment doses) and proceeding to cardioversion rather than delaying cardioversion for 3 weeks of therapeutic anticoagulation or a TEE-guided approach (Grade 2C)**. After successful cardioversion to sinus rhythm, we recommend therapeutic anticoagulation for at least 4 weeks rather than no anticoagulation, regardless of baseline stroke risk** (Grade 2C)**. Decisions about long-term anticoagulation after cardioversion should be made in accordance with our risk-based recommendations for long-term antithrombotic therapy in section 2.1.**

4.2 Patients Undergoing Urgent Cardioversion for Hemodynamically Unstable AF

There are no published data regarding the optimal anticoagulation strategy to use before or during urgent cardioversion for patients with AF and hemodynamic instability. Initiation of anticoagulation immediately before urgent cardioversion (eg, with IV unfractionated heparin or low-molecular-weight heparin) would be expected to reduce the risk of stroke or thromboembolism based on studies

of elective cardioversion. It is important to note that the initiation of anticoagulation therapy should not delay any emergency interventions to stabilize the patient.

Recommendation

4.2. For patients with AF and hemodynamic instability undergoing urgent cardioversion (electrical or pharmacologic), we suggest that therapeutic-dose parenteral anticoagulation be started before cardioversion, if possible (Grade 2C)**, but that initiation of anticoagulation must not delay any emergency intervention** (Grade 2C)**. After successful cardioversion to sinus rhythm, we suggest therapeutic anticoagulation for at least 4 weeks after successful cardioversion to sinus rhythm rather than no anticoagulation, regardless of baseline stroke risk** (Grade 2C)**. Decisions about anticoagulation beyond 4 weeks should be made in accordance with our risk-based recommendations for long-term antithrombotic therapy in section 2.1.**

4.3 Patients Undergoing Elective or Urgent Cardioversion for Atrial Flutter

There are no published data regarding the optimal anticoagulation strategy to use for patients undergoing cardioversion for atrial flutter. Although some observational studies suggest that the risk of thromboembolism after cardioversion for atrial flutter is low, even without anticoagulation, other studies have documented a similar risk of thromboembolism in patients after cardioversion for atrial flutter and AF.[140,162,163] This may be because AF and atrial flutter often coexist.

Recommendation

4.3. For patients with atrial flutter undergoing elective or urgent pharmacologic or electrical cardioversion, we suggest that the same approach to thromboprophylaxis be used as for patients with atrial fibrillation undergoing cardioversion.

5.0 Practical Issues in the Management of VKA Therapy

5.1 Optimal Target INR Range

For a full discussion of optimal target INR range with VKA therapy across a variety of indications, see Holbrook et al[91] regarding evidence-based management of anticoagulation in this guideline. With respect to patients with AF specifically, several studies assessed oral anticoagulation at very-low INR targets or fixed low doses compared with adjusted-dose anticoagulation targeted at an INR of 2.0 to 3.0 and found that anticoagulation targeted at an INR of 2.0 to 3.0 was more effective in reducing the risk of stroke.[164] Observational studies have shown that the risk of ischemic stroke is much greater once INR levels are <2.0 and that there are no appreciable gains in efficacy with levels >2.0. However, there is a sharp increase in the risk of bleeding complications, particularly intracranial hemorrhage, as INR levels rise to >3.0 to 4.0.[165-169] A target INR value of 2.0 will result in patients spending a substantial proportion of time at subtherapeutic INR levels (ie, <2.0).[170] Given that bleeding risk does not rise substantially until INR levels are >3.0, and particularly >4.0, these data from AF populations support an optimal target INR range of 2.0 to 3.0, with a target value of 2.5 to maximize the time spent in the optimal INR range.[165-170]

5.2 Time in Therapeutic Range

For a full discussion of the importance of time in therapeutic range while on adjusted-dose VKA therapy, see Ageno et al[171] regarding oral anticoagulation in this guideline. With respect to patients with AF, there are several observational studies indicating that increasing time out of range is associated with poorer outcomes (eg, mortality, ischemic stroke, thromboembolism, major bleeding).[172-175]

6.0 Future Research

Approximately one in every three patients with AF also has coronary artery disease.[51] However, the optimal approach to antithrombotic therapy in these patients is unclear. Research is needed to determine the effect of treatment with oral anticoagulation and aspirin compared with oral anticoagulation alone on patient-important outcomes of vascular death, nonfatal stroke, nonfatal MI, nonfatal major extracranial bleeding, and nonfatal systemic embolism. Research is also needed to inform recommendations about different antithrombotic therapy regimens for patients with AF undergoing placement of an intracoronary artery stent or who experience an ACS (eg, triple therapy with oral anticoagulation, clopidogrel, and aspirin; dual antiplatelet therapy with clopidogrel and aspirin; or combination oral anticoagulation and aspirin or clopidogrel). Finally, all existing stroke risk stratification and bleeding risk stratification schema for patients with AF have modest predictive value, and development of more-accurate risk stratification systems is needed to facilitate a more-accurate estimation of net clinical benefit for individual patients.

7.0 Conclusions

Stroke is a serious complication of AF, but its risk varies considerably across different groups of patients with AF. Antithrombotic prophylaxis for stroke is associated with an increased risk of bleeding. We provide recommendations for antithrombotic treatment based on net clinical benefit for patients with AF at varying levels of stroke risk. Oral anticoagulation is the optimal choice of antithrombotic therapy for patients with AF at high risk of stroke (eg, $CHADS_2$ score ≥ 2). At lower levels of stroke risk, antithrombotic treatment decisions will require a more individualized approach that takes into consideration patient values and preferences, bleeding risk, and the presence of non-$CHADS_2$ stroke risk factors. The role of oral anticoagulation for the prevention of stroke in patients with AF will evolve as the results of large, ongoing, phase 3 RCTs of new oral anticoagulants are published and as experience with these new agents in clinical practice continues to grow.

Acknowledgments

Author contributions: As Topic Editor, Dr You oversaw the development of this article, including the data analysis and subsequent development of the recommendations contained herein.
Dr You: contributed as Topic Editor.
Dr Singer: contributed as a panelist.
Dr Howard: contributed as a panelist.
Dr Lane: contributed as a panelist.
Dr Eckman: contributed as a resource consultant.
Dr Fang: contributed as a panelist.
Dr Hylek: contributed as a panelist.
Dr Schulman: contributed as a panelist.
Dr Go: contributed as a panelist.
Dr Hughes: contributed as a panelist.
Dr Spencer: contributed as a panelist.
Dr Manning: contributed as a panelist.
Dr Halperin: contributed as a panelist.
Dr Lip: contributed as Deputy Editor and senior author.

Financial/nonfinancial disclosures: The authors of this guideline provided detailed conflict of interest information related to each individual recommendation made in this article. A grid of these disclosures is available online at http://chestjournal.chestpubs.org/content/141/2_suppl/e531S/suppl/DC1. In summary, the authors have reported to *CHEST* the following conflicts of interest: Over the past 3 years, Dr Singer has received grant support from the Eliot B. and Edith C. Shoolman fund of Massachusetts General Hospital (MGH) to study stroke in atrial fibrillation. He has also received research grant support (through MGH) from Daiichi-Sankyo, Inc, to study hemorrhage risk in patients with atrial fibrillation taking warfarin. Dr Singer has been supported by the National Institutes of Health (National Heart, Lung, and Blood Institute and National Institute on Aging) grant funding to study stroke risk in patients with atrial fibrillation. He has served as a consultant on stroke prevention in atrial fibrillation to Bayer Healthcare Pharmaceuticals; Boehringer Ingelheim GmbH; Bristol-Myers Squibb; Daiichi-Sankyo, Inc; Johnson & Johnson; Merck and Co, Inc; Pfizer Inc; and Sanofi-Aventis LLC. Dr. Singer has published numerous articles on stroke prevention in atrial fibrillation and has given numerous lectures and research presentations on this topic. Dr Lane is in receipt of an investigator-initiated educational grant from Bayer Healthcare Pharmaceuticals and has participated in speaker activities for Bayer Healthcare Pharmaceuticals, Boehringer Ingelheim GmbH, Bristol-Myers Squibb/Pfizer Inc, and the Thrombosis Research Institute. She has also received support (travel and accommodation) to attend conferences from AstraZeneca and Boehringer Ingelheim GmbH. Dr Eckman has received a number of grants, including university grants for "Using Decision Analytic Modeling to Guide the ACCP Guideline Development Process for Antithrombotic Therapy in Atrial Fibrillation" (Foundation for Informed Medical Decision Making; $185,000) and has the following industry grants: "Cost-Effectiveness of Screening for Chronic Hepatitis C Infection" (Merck/Schering-Plough; September 2012; $58,000) and "Cost-Effectiveness of Screening for Chronic Hepatitis B Infection" (Gilead Sciences Inc; $56,000). He served as a consultant for Savient Pharmaceuticals (~$300). Dr Hylek has participated in a symposium sponsored by Boehringer Ingelheim, served on advisory boards (Bayer Healthcare Pharmaceuticals; Boehringer Ingelheim GmbH; Bristol-Myers Squibb; Daiichi-Sankyo, Inc; Johnson & Johnson; Merck and Co, Inc; Ortho-McNeil Pharmaceutical, Inc; and Pfizer Inc) for amounts totaling < $10,000 and participated at the steering committee level for several pharmaceutical-sponsored studies (ARISTOTLE trial sponsored by Bristol-Myers Squibb and Pfizer Inc [< $10,000] and ORBIT-AF Registry sponsored by Ortho-McNeil Pharmaceutical, Inc [< $10,000]). Dr Go has received research funding from the National Heart, Lung, and Blood Institute related to antithrombotic therapy in atrial fibrillation and was a site principal investigator for a clinical trial sponsored by Johnson & Johnson and Bayer Healthcare Pharmaceuticals. Dr Halperin has received consulting fees from the following pharmaceutical manufacturers for advisory activities involving the development of anticoagulant drugs, none of which are currently approved for clinical use in any indication in the United States: Astellas Pharma, US; Bayer AG HealthCare; Boehringer Ingelheim; Daiichi Sankyo; Johnson & Johnson; and Sanofi-Aventis. He has received honoraria from Portola Pharmaceuticals, Inc as a member of the Data Safety Monitoring Board of its Phase II EXPLORE-AF trial involving an investigational anticoagulant for prevention of thromboembolism in patients with atrial fibrillation. He has received consulting fees from Biotronik, Inc as co-chair of the Steering Committee for the IMPACT clinical trial evaluating the use ambulatory monitoring technology in approved implanted cardiac arrhythmia devices to guide anticoagulation therapy for stroke prevention. He has received a consulting fee from the Bristol-Myers Squibb/Sanofi Partnership for advisory activities related to the use of the platelet inhibitor drug, clopidogrel, for prevention of thromboembolism in patients with atrial fibrillation. He has been a speaker at CME Symposia that derived partial funding from the following sponsors involved in the development of anticoagulants for potential use in patients with AF: Bayer AG HealthCare, Boehringer Ingelheim, and Sanofi-Aventis. Dr Lip has served as a consultant for Bayer, Astellas, Merck, Daiichi-Sankyo, AstraZeneca, Sanofi-Aventis, BMS/Pfizer, Biotronik, Portola, and Boehringer Ingelheim and has been on the speakers bureau for Bayer, Bristol-Myers Squibb/Pfizer, Boehringer Ingelheim, and Sanofi-Aventis. Drs You, Howard, Fang, Schulman, Hughes, Manning, and Spencer have reported that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of sponsors: The sponsors played no role in the development of these guidelines. Sponsoring organizations cannot recommend panelists or topics, nor are they allowed prepublication access to the manuscripts and recommendations. Guideline panel members, including the chair, and members of the Health & Science Policy Committee are blinded to the funding sources. Further details on the Conflict of Interest Policy are available online at http://chestnet.org.

Other contributions: We thank Louis Kuritzky, MD, for sharing his perspective as a frontline clinician during the development of our recommendations.

Endorsements: This guideline is endorsed by the American Association for Clinical Chemistry, the American College of Clinical Pharmacy, the American Society of Health-System Pharmacists, the American Society of Hematology, and the International Society of Thrombosis and Hematosis.

Additional information: The supplement Tables can be found in the Online Data Supplement at http://chestjournal.chestpubs.org/content/141/2_suppl/e531S/suppl/DC1.

References

1. Lloyd-Jones DM, Wang TJ, Leip EP, et al. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. *Circulation*. 2004;110(9):1042-1046.
2. Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation*. 2006;114(2):119-125.
3. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke*. 1991;22(8):983-988.
4. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials [published correction in *Arch Intern Med*. 1994;154(19):2254]. *Arch Intern Med*. 1994; 154(13):1449-1457.
5. Douketis JD, Spyropoulos AC, Spencer FA, et al. Perioperative management of antithrombotic therapy: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e326S-e350S.
6. Lansberg MG, O'Donnell MJ, Khatri P, et al. Antithrombotic and thrombolytic therapy for ischemic stroke: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e601S-e636S.
7. Whitlock RP, Sun JC, Fremes SE, Rubens FD, Teoh KH. Antithrombotic and thrombolytic therapy for valvular disease: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl): e576S-e600S.
8. Bates SM, Greer IA, Middeldorp S, Veenstra D, Prabulos A-M, Vandvik PO. VTE, thrombophilia, antithrombotic therapy, and pregnancy: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e691S-e736S.
9. MacLean S, Mulla S, Akl EA, et al. Patient values and preferences in decision making for antithrombotic therapy: a systematic review: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e1S-e23S.
10. Guyatt GH, Norris SL, Schulman S, et al. Methodology for the development of antithrombotic therapy and prevention of thrombosis guidelines: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):53S-70S.
11. Andersen KK, Olsen TS, Dehlendorff C, Kammersgaard LP. Hemorrhagic and ischemic strokes compared: stroke severity, mortality, and risk factors. *Stroke*. 2009;40(6):2068-2072.
12. Stroke Risk in Atrial Fibrillation Working Group. Comparison of 12 risk stratification schemes to predict stroke in patients with nonvalvular atrial fibrillation. *Stroke*. 2008; 39(6):1901-1910.
13. Lévy S, Camm AJ, Saksena S, et al; Working Group on Arrhythmias, Working Group on Cardiac Pacing of the European Society of Cardiology, North American Society of Pacing and Electrophysiology. International consensus on nomenclature and classification of atrial fibrillation; a collaborative project of the Working Group on Arrhythmias and the Working Group on Cardiac Pacing of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Europace*. 2003;5(2): 119-122.
14. Petersen P. Thromboembolic complications in atrial fibrillation. *Stroke*. 1990;21(1):4-13.
15. Brand FN, Abbott RD, Kannel WB, Wolf PA. Characteristics and prognosis of lone atrial fibrillation. 30-year follow-up in the Framingham Study. *JAMA*. 1985;254(24):3449-3453.
16. Hart RG, Pearce LA, Rothbart RM, McAnulty JH, Asinger RW, Halperin JL; Stroke Prevention in Atrial Fibrillation Investigators. Stroke with intermittent atrial fibrillation: incidence and predictors during aspirin therapy. *J Am Coll Cardiol*. 2000;35(1):183-187.
17. Hohnloser SH, Pajitnev D, Pogue J, et al; ACTIVE W Investigators. Incidence of stroke in paroxysmal versus sustained atrial fibrillation in patients taking oral anticoagulation or combined antiplatelet therapy: an ACTIVE W Substudy. *J Am Coll Cardiol*. 2007;50(22):2156-2161.
18. Stroke Risk in Atrial Fibrillation Working Group. Independent predictors of stroke in patients with atrial fibrillation: a systematic review. *Neurology*. 2007;69(6):546-554.
19. Hughes M, Lip GY; Guideline Development Group, National Clinical Guideline for Management of Atrial Fibrillation in Primary and Secondary Care, National Institute for Health and Clinical Excellence. Stroke and thromboembolism in atrial fibrillation: a systematic review of stroke risk factors, risk stratification schema and cost effectiveness data. *Thromb Haemost*. 2008;99(2):295-304.
20. van Walraven C, Hart RG, Connolly S, et al. Effect of age on stroke prevention therapy in patients with atrial fibrillation: the atrial fibrillation investigators. *Stroke*. 2009;40(4): 1410-1416.
21. Lin LY, Lee CH, Yu CC, et al. Risk factors and incidence of ischemic stroke in Taiwanese with nonvalvular atrial fibrillation—a nation wide database analysis. *Atherosclerosis*. 2011;217(1):292-295.
22. Frost L, Engholm G, Johnsen S, Møller H, Husted S. Incident stroke after discharge from the hospital with a diagnosis of atrial fibrillation. *Am J Med*. 2000;108(1):36-40.
23. The Stroke Prevention in Atrial Fibrillation Investigators Committee on Echocardiography. Transesophageal echocardiographic correlates of thromboembolism in high-risk patients with nonvalvular atrial fibrillation. *Ann Intern Med*. 1998;128(8):639-647.
24. Olesen JB, Lip GY, Hansen ML, et al. Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ*. 2011;342:d124.
25. Stroke Prevention in Atrial Fibrillation Investigators. Risk factors for thromboembolism during aspirin therapy in patients with atrial fibrillation: the Stroke Prevention in Atrial Fibrillation Study. *J Stroke Cerebrovasc Dis*. 1995;5(3): 147-157.
26. van Latum JC, Koudstaal PJ, Venables GS, van Gijn J, Kappelle LJ, Algra A; European Atrial Fibrillation Trial (EAFT) Study Group. Predictors of major vascular events in patients with a transient ischemic attack or minor ischemic stroke and with nonrheumatic atrial fibrillation. *Stroke*. 1995;26(5):801-806.
27. Echocardiographic predictors of stroke in patients with atrial fibrillation: a prospective study of 1066 patients from 3 clinical trials. *Arch Intern Med*. 1998;158(12):1316-1320.
28. Hart RG, Pearce LA, McBride R, Rothbart RM, Asinger RW; The Stroke Prevention in Atrial Fibrillation (SPAF) Investigators. Factors associated with ischemic stroke during aspirin therapy in atrial fibrillation: analysis of 2012 participants in the SPAF I-III clinical trials. *Stroke*. 1999;30(6):1223-1229.
29. Fuster V, Rydén LE, Cannom DS, et al; American College of Cardiology/American Heart Association Task Force

on Practice Guidelines; European Society of Cardiology Committee for Practice Guidelines; European Heart Rhythm Association; Heart Rhythm Society. ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation): developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Circulation*. 2006;114(7):e257-e354.
30. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA*. 2001;285(22):2864-2870.
31. Albers GW, Dalen JE, Laupacis A, Manning WJ, Petersen P, Singer DE. Antithrombotic therapy in atrial fibrillation. *Chest*. 2001;119(suppl 1):194S-206S.
32. Wang TJ, Massaro JM, Levy D, et al. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA*. 2003;290(8):1049-1056.
33. van Walraven C, Hart RG, Wells GA, et al. A clinical prediction rule to identify patients with atrial fibrillation and a low risk for stroke while taking aspirin. *Arch Intern Med*. 2003;163(8):936-943.
34. Singer DE, Albers GW, Dalen JE, Go AS, Halperin JL, Manning WJ. Antithrombotic therapy in atrial fibrillation: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest*. 2004;126(suppl 3):429S-456S.
35. Lip GY, Lane D, Van Walraven C, Hart RG. Additive role of plasma von Willebrand factor levels to clinical factors for risk stratification of patients with atrial fibrillation. *Stroke*. 2006;37(9):2294-2300.
36. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest*. 2010;137(2):263-272.
37. Gage BF, van Walraven C, Pearce L, et al. Selecting patients with atrial fibrillation for anticoagulation: stroke risk stratification in patients taking aspirin. *Circulation*. 2004;110(16):2287-2292.
38. Baruch L, Gage BF, Horrow J, et al. Can patients at elevated risk of stroke treated with anticoagulants be further risk stratified? *Stroke*. 2007;38(9):2459-2463.
39. Fang MC, Go AS, Chang Y, Borowsky L, Pomernacki NK, Singer DE; ATRIA Study Group. Comparison of risk stratification schemes to predict thromboembolism in people with nonvalvular atrial fibrillation. *J Am Coll Cardiol*. 2008;51(8):810-815.
40. Poli D, Lip GY, Antonucci E, Grifoni E, Lane D. Stroke risk stratification in a "real-world" elderly anticoagulated atrial fibrillation population. *J Cardiovasc Electrophysiol*. 2010;22(1):25-30.
41. Ruiz Ortiz M, Romo E, Mesa D, et al. Oral anticoagulation in nonvalvular atrial fibrillation in clinical practice: impact of CHADS(2) score on outcome. *Cardiology*. 2010;115(3):200-204.
42. Lip GY, Frison L, Halperin JL, Lane DA. Identifying patients at high risk for stroke despite anticoagulation: a comparison of contemporary stroke risk stratification schemes in an anticoagulated atrial fibrillation cohort. *Stroke*. 2010;41(12):2731-2738.
43. Van Staa TP, Setakis E, Di Tanna GL, Lane DA, Lip GY. A comparison of risk stratification schemes for stroke in 79,884 atrial fibrillation patients in general practice. *J Thromb Haemost*. 2011;9(1):39-48.
44. Go AS, Hylek EM, Chang Y, et al. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? *JAMA*. 2003;290(20):2685-2692.
45. Fang MC, Singer DE, Chang Y, et al. Gender differences in the risk of ischemic stroke and peripheral embolism in atrial fibrillation: the AnTicoagulation and Risk factors In Atrial fibrillation (ATRIA) study. *Circulation*. 2005;112(12):1687-1691.
46. Healey JS, Hart RG, Pogue J, et al. Risks and benefits of oral anticoagulation compared with clopidogrel plus aspirin in patients with atrial fibrillation according to stroke risk: the atrial fibrillation clopidogrel trial with irbesartan for prevention of vascular events (ACTIVE-W). *Stroke*. 2008;39(5):1482-1486.
47. Arima H, Hart RG, Colman S, et al; PROGRESS Collaborative Group. Perindopril-based blood pressure-lowering reduces major vascular events in patients with atrial fibrillation and prior stroke or transient ischemic attack. *Stroke*. 2005;36(10):2164-2169.
48. Lip GY, Frison L, Grind M; SPORTIF Investigators. Effect of hypertension on anticoagulated patients with atrial fibrillation. *Eur Heart J*. 2007;28(6):752-759.
49. Eckman MH, Singer DE, Rosand J, Greenberg SM. Moving the tipping point: the decision to anticoagulate patients with atrial fibrillation. *Circ Cardiovasc Qual Outcomes*. 2011;4(1):14-21.
50. National Collaborating Centre for Chronic Conditions. Atrial fibrillation: national clinical guideline for management in primary and secondary care. London, England: Royal College of Physicians; 2006:104. Available at: http://www.nice.org.uk/nicemedia/live/10982/30055/30055.pdf. Accessed November 22, 2011.
51. Singer DE, Chang Y, Fang MC, et al. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. *Ann Intern Med*. 2009;151(5):297-305.
52. Hart RG, Pearce LA. Current status of stroke risk stratification in patients with atrial fibrillation. *Stroke*. 2009;40(7):2607-2610.
53. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: The Euro Heart Survey. *Chest*. 2010;138(5):1093-1100.
54. van der Meer FJ, Rosendaal FR, Vandenbroucke JP, Briët E. Bleeding complications in oral anticoagulant therapy. An analysis of risk factors. *Arch Intern Med*. 1993;153(13):1557-1562.
55. Petersen P, Boysen G, Godtfredsen J, Andersen ED, Andersen B. Placebo-controlled, randomised trial of warfarin and aspirin for prevention of thromboembolic complications in chronic atrial fibrillation. The Copenhagen AFASAK study. *Lancet*. 1989;1(8631):175-179.
56. The Boston Area Anticoagulation Trial for Atrial Fibrillation Investigators. The effect of low-dose warfarin on the risk of stroke in patients with nonrheumatic atrial fibrillation. *N Engl J Med*. 1990;323(22):1505-1511.
57. Connolly SJ, Laupacis A, Gent M, Roberts RS, Cairns JA, Joyner C. Canadian Atrial Fibrillation Anticoagulation (CAFA) Study. *J Am Coll Cardiol*. 1991;18(2):349-355.
58. Stroke Prevention in Atrial Fibrillation Study. Final results. *Circulation*. 1991;84(2):527-539.
59. Ezekowitz MD, Bridgers SL, James KE, et al; Veterans Affairs Stroke Prevention in Nonrheumatic Atrial Fibrillation Investigators. Warfarin in the prevention of stroke associated with nonrheumatic atrial fibrillation. *N Engl J Med*. 1992;327(20):1406-1412.

60. EAFT (European Atrial Fibrillation Trial) Study Group. Secondary prevention in non-rheumatic atrial fibrillation after transient ischaemic attack or minor stroke. *Lancet*. 1993;342(8882):1255-1262.
61. van Walraven C, Hart RG, Singer DE, et al. Oral anticoagulants vs aspirin in nonvalvular atrial fibrillation: an individual patient meta-analysis. *JAMA*. 2002;288(19):2441-2448
62. Benavente O, Hart R, Koudstaal P, Laupacis A, McBride R. Antiplatelet therapy for preventing stroke in patients with atrial fibrillation and no previous history of stroke or transient ischemic attacks. In: Warlow C, Van Gijn J, Sandercock P, eds. *Stroke Module of the Cochrane Database of Systematic Reviews*. Oxford, England: The Cochrane Collaboration; 1999.
63. Posada IS, Barriales V; LASAF Pilot Study Group. Alternate-day dosing of aspirin in atrial fibrillation. *Am Heart J*. 1999; 138(1 pt 1):137-143.
64. Sato H, Ishikawa K, Kitabatake A, et al; Japan Atrial Fibrillation Stroke Trial Group. Low-dose aspirin for prevention of stroke in low-risk patients with atrial fibrillation: Japan Atrial Fibrillation Stroke Trial. *Stroke*. 2006;37(2): 447-451.
65. Diener HC, Lowenthal A. Antiplatelet therapy to prevent stroke: risk of brain hemorrhage and efficacy in atrial fibrillation. *J Neurol Sci*. 1997;153(1):112.
66. Edvardsson N, Juul-Möller S, Omblus R, Pehrsson K. Effects of low-dose warfarin and aspirin versus no treatment on stroke in a medium-risk patient population with atrial fibrillation. *J Intern Med*. 2003;254(1):95-101.
67. Baigent C, Blackwell L, Collins R, et al; Antithrombotic Trialists' (ATT) Collaboration. Aspirin in the primary and secondary prevention of vascular disease: collaborative meta-analysis of individual participant data from randomised trials. *Lancet*. 2009;373(9678):1849-1860.
68. Antithrombotic Trialists' Collaboration; Baigent C, Blackwell L, Collins R, et al. Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002;324(7329):71-86.
69. Vemmos KN, Tsivgoulis G, Spengos K, et al. Primary prevention of arterial thromboembolism in the oldest old with atrial fibrillation—a randomized pilot trial comparing adjusted-dose and fixed low-dose coumadin with aspirin. *Eur J Intern Med*. 2006;17(1):48-52.
70. Gulløv AL, Koefoed BG, Petersen P, et al. Fixed minidose warfarin and aspirin alone and in combination vs adjusted-dose warfarin for stroke prevention in atrial fibrillation: Second Copenhagen Atrial Fibrillation, Aspirin, and Anticoagulation Study. *Arch Intern Med*. 1998;158(14):1513-1521.
71. Mant J, Hobbs FD, Fletcher K, et al; BAFTA Investigators; Midland Research Practices Network (MidReC). Warfarin versus aspirin for stroke prevention in an elderly community population with atrial fibrillation (the Birmingham Atrial Fibrillation Treatment of the Aged Study, BAFTA): a randomised controlled trial. *Lancet*. 2007;370(9586):493-503.
72. Pérez-Gómez F, Alegría E, Berjón J, et al; NASPEAF Investigators. Comparative effects of antiplatelet, anticoagulant, or combined therapy in patients with valvular and nonvalvular atrial fibrillation: a randomized multicenter study. *J Am Coll Cardiol*. 2004;44(8):1557-1566.
73. Hellemons BS, Langenberg M, Lodder J, et al. Primary prevention of arterial thromboembolism in non-rheumatic atrial fibrillation in primary care: randomised controlled trial comparing two intensities of coumarin with aspirin. *BMJ*. 1999;319(7215):958-964.
74. Morocutti C, Amabile G, Fattapposta F, et al; SIFA (Studio Italiano Fibrillazione Atriale) Investigators. Indobufen versus warfarin in the secondary prevention of major vascular events in nonrheumatic atrial fibrillation. *Stroke*. 1997; 28(5):1015-1021.
75. Warfarin versus aspirin for prevention of thromboembolism in atrial fibrillation: Stroke Prevention in Atrial Fibrillation II Study. *Lancet*. 1994;343(8899):687-691.
76. Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation: Stroke Prevention in Atrial Fibrillation III randomised clinical trial. *Lancet*. 1996;348(9028):633-638.
77. Rash A, Downes T, Portner R, Yeo WW, Morgan N, Channer KS. A randomised controlled trial of warfarin versus aspirin for stroke prevention in octogenarians with atrial fibrillation (WASPO). *Age Ageing*. 2007;36(2):151-156.
78. Anand SS, Yusuf S. Oral anticoagulants in patients with coronary artery disease. *J Am Coll Cardiol*. 2003;41(4)(suppl S): 62S-69S.
79. Chimowitz MI, Lynn MJ, Howlett-Smith H, et al; Warfarin-Aspirin Symptomatic Intracranial Disease Trial Investigators. Comparison of warfarin and aspirin for symptomatic intracranial arterial stenosis. *N Engl J Med*. 2005;352(13):1305-1316.
80. Connolly S, Pogue J, Hart R, et al; ACTIVE Writing Group of the ACTIVE Investigators. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation in the Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE W): a randomised controlled trial. *Lancet*. 2006;367(9526):1903-1912.
81. Connolly SJ, Pogue J, Hart RG, et al; ACTIVE Investigators. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. *N Engl J Med*. 2009;360(20):2066-2078.
82. Olsson SB; Executive Steering Committee of the SPORTIF III Investigators. Stroke prevention with the oral direct thrombin inhibitor ximelagatran compared with warfarin in patients with non-valvular atrial fibrillation (SPORTIF III): randomised controlled trial. *Lancet*. 2003; 362(9397):1691-1698.
83. Albers GW, Diener HC, Frison L, et al; SPORTIF Executive Steering Committee for the SPORTIF V Investigators. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation: a randomized trial. *JAMA*. 2005;293(6):690-698.
84. Connolly SJ, Ezekowitz MD, Yusuf S, et al; RE-LY Steering Committee and Investigators. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med*. 2009;361(12):1139-1151.
85. Patel MR, Mahaffey KW, Garg J, et al; ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med*. 2011;365(10):883-891.
86. Connolly SJ, Eikelboom J, Joyner C, et al; AVERROES Steering Committee and Investigators. Apixaban in patients with atrial fibrillation. *N Engl J Med*. 2011;364(9):806-817.
87. Granger CB, Alexander JH, McMurray JJ, et al; ARISTOTLE Committees and Investigators. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med*. 2011;365(11): 981-992.
88. Roskell NS, Lip GY, Noack H, Clemens A, Plumb JM. Treatments for stroke prevention in atrial fibrillation: a network meta-analysis and indirect comparisons versus dabigatran etexilate. *Thromb Haemost*. 2010;104(6):1106-1115.
89. Label approved on 03/04/2011 for Pradaxa, NDA no. 022512. In: United States Food and Drug Administration; 2010.
90. van Ryn J, Stangier J, Haertter S, et al. Dabigatran etexilate—a novel, reversible, oral direct thrombin inhibitor: interpretation of coagulation assays and reversal of anticoagulant activity. *Thromb Haemost*. 2010;103(6):1116-1127.
91. Holbrook A, Schulman S, Witt DM, et al. Evidence-based management of anticoagulant therapy: antithrombotic

therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e152S-e184S.
92. Bloomfield HE, Krause A, Greer N, et al. Meta-analysis: effect of patient self-testing and self-management of long-term anticoagulation on major clinical outcomes. *Ann Intern Med*. 2011;154(7):472-482.
93. Freeman JV, Zhu RP, Owens DK, et al. Cost-effectiveness of dabigatran compared with warfarin for stroke prevention in atrial fibrillation. *Ann Intern Med*. 2011;154(1):1-11.
94. Hughes M, Lip GY; Guideline Development Group for the NICE national clinical guideline for management of atrial fibrillation in primary and secondary care. Risk factors for anticoagulation-related bleeding complications in patients with atrial fibrillation: a systematic review. *QJM*. 2007;100(10):599-607.
95. Beyth RJ, Quinn LM, Landefeld CS. Prospective evaluation of an index for predicting the risk of major bleeding in outpatients treated with warfarin. *Am J Med*. 1998;105(2):91-99.
96. Landefeld CS, Goldman L. Major bleeding in outpatients treated with warfarin: incidence and prediction by factors known at the start of outpatient therapy. *Am J Med*. 1989;87(2):144-152.
97. Shireman TI, Mahnken JD, Howard PA, Kresowik TF, Hou Q, Ellerbeck EF. Development of a contemporary bleeding risk model for elderly warfarin recipients. *Chest*. 2006;130(5):1390-1396.
98. Gage BF, Yan Y, Milligan PE, et al. Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). *Am Heart J*. 2006;151(3):713-719.
99. Lip GY, Frison L, Halperin JL, Lane DA. Comparative validation of a novel risk score for predicting bleeding risk in anticoagulated patients with atrial fibrillation: the HAS-BLED (Hypertension, Abnormal Renal/Liver Function, Stroke, Bleeding History or Predisposition, Labile INR, Elderly, Drugs/Alcohol Concomitantly) score. *J Am Coll Cardiol*. 2011;57(2):173-180.
100. Holmes DR, Reddy VY, Turi ZG, et al; PROTECT AF Investigators. Percutaneous closure of the left atrial appendage versus warfarin therapy for prevention of stroke in patients with atrial fibrillation: a randomised non-inferiority trial. *Lancet*. 2009;374(9689):534-542.
101. Ostermayer SH, Reisman M, Kramer PH, et al. Percutaneous left atrial appendage transcatheter occlusion (PLAATO system) to prevent stroke in high-risk patients with non-rheumatic atrial fibrillation: results from the international multi-center feasibility trials. *J Am Coll Cardiol*. 2005;46(1):9-14.
102. Block PC, Burstein S, Casale PN, et al. Percutaneous left atrial appendage occlusion for patients in atrial fibrillation suboptimal for warfarin therapy: 5-year results of the PLAATO (Percutaneous Left Atrial Appendage Transcatheter Occlusion) Study. *JACC Cardiovasc Interv*. 2009;2(7):594-600.
103. Lechat P, Lardoux H, Mallet A, et al; FFAACS (Fluindione, Fibrillation Auriculaire, Aspirin et Contraste Spontané) Investigators. Anticoagulant (fluindione)-aspirin combination in patients with high-risk atrial fibrillation. A randomized trial (Fluindione, Fibrillation Auriculaire, Aspirin et Contraste Spontané; FFAACS). *Cerebrovasc Dis*. 2001;12(3):245-252.
104. Flaker GC, Gruber M, Connolly SJ, et al; SPORTIF Investigators. Risks and benefits of combining aspirin with anticoagulant therapy in patients with atrial fibrillation: an exploratory analysis of stroke prevention using an oral thrombin inhibitor in atrial fibrillation (SPORTIF) trials. *Am Heart J*. 2006;152(5):967-973.
105. Clinical review for Pradaxa, NDA 22-512. In: United States Food and Drug Administration; 2010:96.
106. Hansen ML, Sørensen R, Clausen MT, et al. Risk of bleeding with single, dual, or triple therapy with warfarin, aspirin, and clopidogrel in patients with atrial fibrillation. *Arch Intern Med*. 2010;170(16):1433-1441.
107. Dentali F, Douketis JD, Lim W, Crowther M. Combined aspirin-oral anticoagulant therapy compared with oral anticoagulant therapy alone among patients at risk for cardiovascular disease: a meta-analysis of randomized trials. *Arch Intern Med*. 2007;167(2):117-124.
108. Vandvik PO, Lincoff AM, Gore JM, et al. Primary and secondary prevention of cardiovascular disease: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e637S-e668S.
109. Iakovou I, Schmidt T, Bonizzoni E, et al. Incidence, predictors, and outcome of thrombosis after successful implantation of drug-eluting stents. *JAMA*. 2005;293(17):2126-2130.
110. Serruys PW, Strauss BH, Beatt KJ, et al. Angiographic follow-up after placement of a self-expanding coronary-artery stent. *N Engl J Med*. 1991;324(1):13-17.
111. Roubin GS, Cannon AD, Agrawal SK, et al. Intracoronary stenting for acute and threatened closure complicating percutaneous transluminal coronary angioplasty. *Circulation*. 1992;85(3):916-927.
112. Foley JB, Brown RI, Penn IM. Thrombosis and restenosis after stenting in failed angioplasty: comparison with elective stenting. *Am Heart J*. 1994;128(1):12-20.
113. Cook S, Windecker S. Early stent thrombosis: past, present, and future. *Circulation*. 2009;119(5):657-659.
114. Cosmi B, Rubboli A, Castelvetri C, Milandri M. Ticlopidine versus oral anticoagulation for coronary stenting. *Cochrane Database Syst Rev*. 2001;(4):CD002133.
115. Lip GY, Huber K, Andreotti F, et al; European Society of Cardiology Working Group on Thrombosis. Management of antithrombotic therapy in atrial fibrillation patients presenting with acute coronary syndrome and/or undergoing percutaneous coronary intervention/ stenting published correction in *Thromb Haemost*. 2010;104(3):653. Dosage error in article text]. *Thromb Haemost*. 2010;103(1):13-28.
116. Faxon DP, Eikelboom JW, Berger PB, Holmes DR, Bhatt DL, Moliterno DJ, Becker RC, Angiolillo DJ. Consensus document: antithrombotic therapy in patients with atrial fibrillation undergoing coronary stenting: a North-American perspective. *Thromb Haemost*. 2011;106(4):572-584.
117. Rossini R, Musumeci G, Lettieri C, et al. Long-term outcomes in patients undergoing coronary stenting on dual oral antiplatelet treatment requiring oral anticoagulant therapy. *Am J Cardiol*. 2008;102(12):1618-1623.
118. Sarafoff N, Ndrepepa G, Mehilli J, et al. Aspirin and clopidogrel with or without phenprocoumon after drug eluting coronary stent placement in patients on chronic oral anticoagulation. *J Intern Med*. 2008;264(5):472-480.
119. Sambola A, Ferreira-González I, Angel J, et al. Therapeutic strategies after coronary stenting in chronically anticoagulated patients: the MUSICA study. *Heart*. 2009;95(18):1483-1488.
120. Gao F, Zhou YJ, Wang ZJ, et al. Comparison of different antithrombotic regimens for patients with atrial fibrillation undergoing drug-eluting stent implantation. *Circ J*. 2010;74(4):701-708.
121. Rothberg MB, Celestin C, Fiore LD, Lawler E, Cook JR. Warfarin plus aspirin after myocardial infarction or the acute coronary syndrome: meta-analysis with estimates of risk and benefit. *Ann Intern Med*. 2005;143(4):241-250.

122. Wallentin L, Becker RC, Budaj A, et al; PLATO Investigators. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med*. 2009;361(11):1045-1057.
123. Holmes DR Jr, Kereiakes DJ, Garg S, et al. Stent thrombosis. *J Am Coll Cardiol*. 2010;56(17):1357-1365.
124. Yusuf S, Zhao F, Mehta SR, Chrolavicius S, Tognoni G, Fox KK. Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. *N Engl J Med*. 2001;345:494-502.
125. Sherman DG, Kim SG, Boop BS, et al; National Heart, Lung, and Blood Institute AFFIRM Investigators. Occurrence and characteristics of stroke events in the Atrial Fibrillation Follow-up Investigation of Sinus Rhythm Management (AFFIRM) study. *Arch Intern Med*. 2005; 165(10):1185-1191.
126. Krittayaphong R, Raungrattanaamporn O, Bhuripanyo K, et al. A randomized clinical trial of the efficacy of radiofrequency catheter ablation and amiodarone in the treatment of symptomatic atrial fibrillation. *J Med Assoc Thai*. 2003;86(suppl 1):S8-S16.
127. Wazni OM, Marrouche NF, Martin DO, et al. Radiofrequency ablation vs antiarrhythmic drugs as first-line treatment of symptomatic atrial fibrillation: a randomized trial. *JAMA*. 2005;293(21):2634-2640.
128. Oral H, Pappone C, Chugh A, et al. Circumferential pulmonary-vein ablation for chronic atrial fibrillation. *N Engl J Med*. 2006;354(9):934-941.
129. Pappone C, Augello G, Sala S, et al. A randomized trial of circumferential pulmonary vein ablation versus antiarrhythmic drug therapy in paroxysmal atrial fibrillation: the APAF Study. *J Am Coll Cardiol*. 2006;48(11):2340-2347.
130. Stabile G, Bertaglia E, Senatore G, et al. Catheter ablation treatment in patients with drug-refractory atrial fibrillation: a prospective, multi-centre, randomized, controlled study (Catheter Ablation For The Cure Of Atrial Fibrillation Study). *Eur Heart J*. 2006;27(2):216-221.
131. Jaïs P, Cauchemez B, Macle L, et al. Catheter ablation versus antiarrhythmic drugs for atrial fibrillation: the A4 study. *Circulation*. 2008;118(24):2498-2505.
132. Wilber DJ, Pappone C, Neuzil P, et al; ThermoCool AF Trial Investigators. Comparison of antiarrhythmic drug therapy and radiofrequency catheter ablation in patients with paroxysmal atrial fibrillation: a randomized controlled trial. *JAMA*. 2010;303(4):333-340.
133. Omran H, Jung W, Rabahieh R, et al. Left atrial appendage function in patients with atrial flutter. *Heart*. 1997;78(3): 250-254.
134. Seidl K, Hauer B, Schwick NG, Zellner D, Zahn R, Senges J. Risk of thromboembolic events in patients with atrial flutter. *Am J Cardiol*. 1998;82(5):580-583.
135. Corrado G, Sgalambro A, Mantero A, et al. Thromboembolic risk in atrial flutter. The FLASIEC (FLutter Atriale Società Italiana di Ecografia Cardiovascolare) multicentre study. *Eur Heart J*. 2001;22(12):1042-1051.
136. Schmidt H, von der Recke G, Illien S, et al. Prevalence of left atrial chamber and appendage thrombi in patients with atrial flutter and its clinical significance. *J Am Coll Cardiol*. 2001;38(3):778-784.
137. Weiss R, Marcovitz P, Knight BP, et al. Acute changes in spontaneous echo contrast and atrial function after cardioversion of persistent atrial flutter. *Am J Cardiol*. 1998; 82(9):1052-1055.
138. Bikkina M, Alpert MA, Mulekar M, Shakoor A, Massey CV, Covin FA. Prevalence of intraatrial thrombus in patients with atrial flutter. *Am J Cardiol*. 1995;76(3):186-189.
139. Moreyra E, Finkelhor RS, Cebul RD. Limitations of transesophageal echocardiography in the risk assessment of patients before nonanticoagulated cardioversion from atrial fibrillation and flutter: an analysis of pooled trials. *Am Heart J*. 1995;129(1):71-75.
140. Gallagher MM, Hennessy BJ, Edvardsson N, et al. Embolic complications of direct current cardioversion of atrial arrhythmias: association with low intensity of anticoagulation at the time of cardioversion. *J Am Coll Cardiol*. 2002; 40(5):926-933.
141. Berger M, Schweitzer P. Timing of thromboembolic events after electrical cardioversion of atrial fibrillation or flutter: a retrospective analysis. *Am J Cardiol*. 1998;82: 1545-1547, A8.
142. Manning WJ, Leeman DE, Gotch PJ, Come PC. Pulsed Doppler evaluation of atrial mechanical function after electrical cardioversion of atrial fibrillation. *J Am Coll Cardiol*. 1989;13(3):617-623.
143. Nagarakanti R, Ezekowitz MD, Oldgren J, et al. Dabigatran versus warfarin in patients with atrial fibrillation: an analysis of patients undergoing cardioversion. *Circulation*. 2011; 123(2):131-136.
144. Manning WJ, Silverman DI, Gordon SP, Krumholz HM, Douglas PS. Cardioversion from atrial fibrillation without prolonged anticoagulation with use of transesophageal echocardiography to exclude the presence of atrial thrombi. *N Engl J Med*. 1993;328(11):750-755.
145. Klein AL, Grimm RA, Murray RD, et al; Assessment of Cardioversion Using Transesophageal Echocardiography Investigators. Use of transesophageal echocardiography to guide cardioversion in patients with atrial fibrillation. *N Engl J Med*. 2001;344(19):1411-1420.
146. Hoppensteadt D, Fareed J, Klein AL, et al; Assessment of Cardioversion Using Transesophageal Echocardiography (ACUTE) II Investigators. Comparison of anticoagulant and anti-inflammatory responses using enoxaparin versus unfractionated heparin for transesophageal echocardiography-guided cardioversion of atrial fibrillation. *Am J Cardiol*. 2008;102(7):842-846.
147. Klein AL, Jasper SE, Katz WE, et al; ACUTE II Steering and Publications Committee for the ACUTE II Investigators. The use of enoxaparin compared with unfractionated heparin for short-term antithrombotic therapy in atrial fibrillation patients undergoing transoesophageal echocardiography-guided cardioversion: Assessment of Cardioversion Using Transoesophageal Echocardiography (ACUTE) II randomized multicentre study. *Eur Heart J*. 2006;27(23): 2858-2865.
148. de Luca I, Sorino M, De Luca L, Colonna P, Del Salvatore B, Corlianò L. Pre- and post-cardioversion transesophageal echocardiography for brief anticoagulation therapy with enoxaparin in atrial fibrillation patients: a prospective study with a 1-year follow-up. *Int J Cardiol*. 2005;102(3): 447-454.
149. Wu LA, Chandrasekaran K, Friedman PA, et al. Safety of expedited anticoagulation in patients undergoing transesophageal echocardiographic-guided cardioversion. *Am J Med*. 2006;119(2):142-146.
150. Stellbrink C, Nixdorff U, Hofmann T, et al; ACE (Anticoagulation in Cardioversion using Enoxaparin) Study Group. Safety and efficacy of enoxaparin compared with unfractionated heparin and oral anticoagulants for prevention of thromboembolic complications in cardioversion of nonvalvular atrial fibrillation: the Anticoagulation in Cardioversion using Enoxaparin (ACE) trial. *Circulation*. 2004;109(8):997-1003.
151. Stangier J. Clinical pharmacokinetics and pharmacodynamics of the oral direct thrombin inhibitor dabigatran etexilate. *Clin Pharmacokinet*. 2008;47(5):285-295.

152. Antonielli E, Pizzuti A, Pálinkás A, et al. Clinical value of left atrial appendage flow for prediction of long-term sinus rhythm maintenance in patients with nonvalvular atrial fibrillation. *J Am Coll Cardiol*. 2002;39(9):1443-1449.
153. Berry C, Stewart S, Payne EM, McArthur JD, McMurray JJ. Electrical cardioversion for atrial fibrillation: outcomes in "real-life" clinical practice. *Int J Cardiol*. 2001;81(1):29-35.
154. Arnar DO, Danielsen R. Factors predicting maintenance of sinus rhythm after direct current cardioversion of atrial fibrillation and flutter: a reanalysis with recently acquired data. *Cardiology*. 1996;87(3):181-188.
155. Paraskevaidis IA, Dodouras T, Tsiapras D, Kremastinos DT. Prediction of successful cardioversion and maintenance of sinus rhythm in patients with lone atrial fibrillation. *Chest*. 2005;127(2):488-494.
156. Pérez Y, Duval AM, Carville C, et al. Is left atrial appendage flow a predictor for outcome of cardioversion of nonvalvular atrial fibrillation? A transthoracic and transesophageal echocardiographic study. *Am Heart J*. 1997;134(4):745-751.
157. Page RL, Wilkinson WE, Clair WK, McCarthy EA, Pritchett EL. Asymptomatic arrhythmias in patients with symptomatic paroxysmal atrial fibrillation and paroxysmal supraventricular tachycardia. *Circulation*. 1994;89(1):224-227.
158. Stoddard MF, Dawkins PR, Prince CR, Ammash NM. Left atrial appendage thrombus is not uncommon in patients with acute atrial fibrillation and a recent embolic event: a transesophageal echocardiographic study. *J Am Coll Cardiol*. 1995;25(2):452-459.
159. Kleemann T, Becker T, Strauss M, Schneider S, Seidl K. Prevalence of left atrial thrombus and dense spontaneous echo contrast in patients with short-term atrial fibrillation <48 hours undergoing cardioversion: value of transesophageal echocardiography to guide cardioversion. *J Am Soc Echocardiogr*. 2009;22(12):1403-1408.
160. Flaker GC, Belew K, Beckman K, et al; AFFIRM Investigators. Asymptomatic atrial fibrillation: demographic features and prognostic information from the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) study. *Am Heart J*. 2005;149(4):657-663.
161. Weigner MJ, Caulfield TA, Danias PG, Silverman DI, Manning WJ. Risk for clinical thromboembolism associated with conversion to sinus rhythm in patients with atrial fibrillation lasting less than 48 hours. *Ann Intern Med*. 1997;126(8):615-620.
162. Arnold AZ, Mick MJ, Mazurek RP, Loop FD, Trohman RG. Role of prophylactic anticoagulation for direct current cardioversion in patients with atrial fibrillation or atrial flutter. *J Am Coll Cardiol*. 1992;19(4):851-855.
163. Chalasani P, Cambre S, Silverman ME. Direct-current cardioversion for the conversion of atrial flutter. *Am J Cardiol*. 1996;77(8):658-660.
164. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. *Ann Intern Med*. 1999;131(7):492-501.
165. Fang MC, Chang Y, Hylek EM, et al. Advanced age, anticoagulation intensity, and risk for intracranial hemorrhage among patients taking warfarin for atrial fibrillation. *Ann Intern Med*. 2004;141(10):745-752.
166. Hylek EM, Go AS, Chang Y, et al. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N Engl J Med*. 2003;349(11):1019-1026.
167. Hylek EM, Skates SJ, Sheehan MA, Singer DE. An analysis of the lowest effective intensity of prophylactic anticoagulation for patients with nonrheumatic atrial fibrillation. *N Engl J Med*. 1996;335(8):540-546.
168. Walker AM, Bennett D. Epidemiology and outcomes in patients with atrial fibrillation in the United States. *Heart Rhythm*. 2008;5(10):1365-1372.
169. Singer DE, Chang Y, Fang MC, et al. Should patient characteristics influence target anticoagulation intensity for stroke prevention in nonvalvular atrial fibrillation?: the ATRIA study. *Circ Cardiovasc Qual Outcomes*. 2009;2(4):297-304.
170. Rose AJ, Ozonoff A, Henault LE, Hylek EM. Warfarin for atrial fibrillation in community-based practise. *J Thromb Haemost*. 2008;6(10):1647-1654.
171. Ageno W, Gallus AS, Wittkowsky A, Crowther M, Hylek EM, Palareti G. Oral anticoagulant therapy: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2)(suppl):e44S-e88S.
172. Jones M, McEwan P, Morgan CL, Peters JR, Goodfellow J, Currie CJ. Evaluation of the pattern of treatment, level of anticoagulation control, and outcome of treatment with warfarin in patients with non-valvar atrial fibrillation: a record linkage study in a large British population. *Heart*. 2005;91(4):472-477.
173. White HD, Gruber M, Feyzi J, et al. Comparison of outcomes among patients randomized to warfarin therapy according to anticoagulant control: results from SPORTIF III and V. *Arch Intern Med*. 2007;167(3):239-245.
174. Connolly SJ, Pogue J, Eikelboom J, et al; ACTIVE W Investigators. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation*. 2008;118(20):2029-2037.
175. Wan Y, Heneghan C, Perera R, et al. Anticoagulation control and prediction of adverse events in patients with atrial fibrillation: a systematic review. *Circ Cardiovasc Qual Outcomes*. 2008;1(2):84-91.

CHEST Online Data Supplement

Antithrombotic Therapy for Atrial Fibrillation

Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines

John J. You, MD; Daniel E. Singer, MD; Patricia A. Howard, PharmD; Deirdre A. Lane, PhD; Mark H. Eckman, MD; Margaret C. Fang, MD, MPH; Elaine M. Hylek, MD, MPH; Sam Schulman, MD, PhD; Alan S. Go, MD; Michael Hughes, PhD; Frederick A. Spencer, MD; Warren J. Manning, MD; Jonathan L. Halperin, MD; and Gregory Y. H. Lip, MD

Table S1—*Studies Assessing Prior Stroke/TIA as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
Pooled estimate[1]	2.5 (1.8-3.5)	.59 (for heterogeneity)
Studies included in pooled estimate:		
AFI[2]/1994	2.5 (NR)	< .05
SPAF I-III Aspirin[3]/1999	2.9 (NR)	< .001
Stöllberger et al[4]/1998	3.7 (1.5-7.5)	.002
Wang et al[5]/2003	1.9 (1.1-3.3)	NR
Studies not included in pooled estimate:		
AFI Echo[6]/1998[a]	3.5 (1.8-6.7)	< .001
Aronow et al[7]/1998	1.6 (1.1-2.2)	< .009
Hart et al[8]/2000 (intermittent AF)	4.1 (NR)	.01
Hart et al[8]/2000 (sustained AF)	2.7 (NR)	< .001
SPAF I[9]/1992	2.1 (NR)	.04
Stöllberger et al[10]/2004	2.14 (NR)	.045
van Latum et al[11]/1995	1.6 (1.0-2.6)	< .05
Van Staa et al[12]/2011	2.86 (2.53-3.22)	< .05

AF = atrial fibrillation; AFI = Atrial Fibrillation Investigators; NR = not reported; RR = risk ratio; SPAF = Stroke Prevention in Atrial Fibrillation; TIA = transient ischemic attack.

[a]Subset of patients from AFI[6] with echocardiographic data available.

 DOI: 10.1378/chest.11-2304

Table S2—*Studies Assessing Hypertension as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
Pooled estimate[1]	2.0 (1.6-2.5)	.29 (for heterogeneity)
Studies included in pooled estimate:		
AFI[2]/1994	1.6 (NR)	< .05
Moulton[13]/1991	1.9 (1.2-3.1)	< .05
SPAF I-III Aspirin[3]/1999	2.0 (NR)	< .001
Stöllberger et al[4]/1998	3.6 (1.8-8.4)	.001
Studies not included in pooled estimate:		
AFI Echo[6]/1998[a]	1.5 (0.9-2.5)	.13
Aronow et al[14]/1989	12.5 (NR)	< .01
Aronow et al[7]/1998	NR	NS
Cabin et al[15]/1990	NR	NS
Hart et al[8]/2000 (intermittent AF)	3.4 (NR)	.003
Hart et al[8]/2000 (sustained AF)	1.8 (NR)	.008
Petersen et al[16]/1990	NR	NS
Seidl et al[17]/1998	6.5 (1.5-4.5)	< .05
SPAF I[9]/1992	2.2 (1.1-4.3)	.02
SPAF III[18]/1998[b]	3.3 (1.7-6.9)	.001
Stöllberger et al[10]/2004	NR	NS
Wang et al[5]/2003[c]	NR	NS
Studies assessing uncontrolled hypertension only[d]:		
Hart et al[8]/2000 (intermittent AF)	NR	NS
Hart et al[8]/2000 (sustained AF)	2.8 (NR)	< .001
SPAF I-III Aspirin[3]/1999	2.3 (NR)	< .001
van Latum et al[11]/1995	1.7 (1.0-2.9)	NS
Van Staa et al[12]/2011 (≥ 180 vs < 120 mm Hg)	2.74 (1.21-6.19)	< .05
Van Staa et al[12]/2011 (140-159 vs < 120 mm Hg)	2.74 (1.21-6.19)	< .05
Van Staa et al[12]/2011 (160-179 vs < 120 mm Hg)	1.49 (0.55-4.00)	NS
Wang et al[5]/2003	1.1/10 mm	< .05

NS = not significant. See Table S1 legend for expansion of other abbreviations.

[a]Subset of patients from AFI (1994)[2] with echocardiographic data available.

[b]Defines hypertension as diagnosed hypertension.

[c]Defines hypertension as use of BP-lowering medication.

[d]Defined as systolic BP > 160 mm Hg, unless otherwise stated.

 DOI: 10.1378/chest.11-2304

Table S3—*Studies Assessing Increasing Age as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
Pooled estimate[1]	1.5 (1.3-1.7)/decade	.59 (for heterogeneity)
Studies included in pooled estimate:		
AFI[2]/1994	1.4 (NR)/decade	< .05
Moulton et al[13]/1991 (for age > 75)	1.7 (1.0-2.8)	< .05
SPAF I-III Aspirin[3]/1999	1.8 (NR)/decade	< .001
Stöllberger et al[4]/1998	1.1 (1.0-1.1)[a]	< .001
Wang et al[5]/2003	1.3 (NR)/decade	< .05
Studies not included in pooled estimate:		
AFI Echo[6]/1998[b]	1.5 (NR)/decade	.006
Cabin et al[15]/1990 (for age > 70)	NR	NS
Hart et al[8]/2000 (intermittent AF)	2.1 (NR)/decade	< .001
Hart et al[8]/2000 (sustained AF)	1.7 (NR)/decade	< .001
Inoue and Atarashi[19]/2000 (paroxysmal AF; for age > 65 y)	3.3 (1.92-5.81)	.0001
Nakagami et al[20]/1998	1.3 (1.0-1.7)/decade	NS
SPAF I[9]/1992	1.2 (0.9-1.6)/decade	NS
SPAF III[18]/1998	1.7 (1.1-2.6)/decade	.01
van Latum et al[11]/1995	NR	NS
Van Staa et al[12]/2011 (age, < 50 vs 60-69 y)	0.14 (0.06-0.34)	< .05
Van Staa et al[12]/2011 (age, ≥ 80 vs 60-69 y)	2.22 (1.78-2.76)	< .05
Van Staa et al[12]/2011 (age, 50-59 vs 60-69 y)	0.44 (0.28-0.69)	< .05
Van Staa et al[12]/2011 (age, 70-79 vs 60-69 y)	1.42 (1.12-1.78)	< .05

See Table S1 and S2 legends for expansion of abbreviations.
[a]Age entered as a continuous variable.
[b]Subset of patients from AFI (1994)[2] with echocardiographic data available.

Table S4—*Studies Assessing Diabetes as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
Pooled estimate[1]	1.7 (1.4-2.0)	.69 (for heterogeneity)
Studies included in pooled estimate:		
AFI[2]/1994	1.7 (NR)	< .05
SPAF I-III Aspirin[3]/1999	1.9 (NR)	.02
Stöllberger et al[4]/1998	NR	NS
Wang et al[5]/2003	1.8 (1.4-3.1)	NR
Studies not included in pooled estimate:		
AFI Echo[6]/1998[a]	1.7 (1.0-2.9)	.05
Aronow et al[7]/1998	NR	NS
Petersen et al[16]/1990	NR	NS
Seidl et al[17]/1998	NR	NS
SPAF III[18]/1998	NR	NS
Van Staa et al[12]/2011	1.33 (1.14-1.55)	< .05

See Table S1 and S2 legends for expansion of abbreviations.
[a]Subset of patients from AFI[2] with echocardiographic data available.

 DOI: 10.1378/chest.11-2304

Table S5—*Studies Assessing Heart Failure as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
AFI[2]/1994	1.4 (NR)	NS
AFI Echo[6]/1998[a] (with echocardiographic data)	1.4 (0.8-2.3)	.16
AFI Echo[6]/1998[a] (without echocardiographic data)	1.7 (1.1-2.7)	.03
SPAF[18]/1998	NR	NS
SPAF I[9]/1992	2.6 (1.0-4.2)	.01
SPAF I-III Aspirin[3]/1999	NR	NS
Stöllberger et al[8]/1998[b]	NR	NS
Stöllberger et al[10]/2004[b]	NR	NS
Van Staa et al[12]/2011	1.26 (1.11-1.42)	< .05

See Table S1 and S2 legends for expansion of abbreviations.
[a]Subset of patients from AFI[2] with echocardiographic data available.
[b]Defined as New York Heart Association functional class > II.

Table S6—*Studies Assessing LV Dysfunction/Hypertrophy (Using Echocardiography) as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
AFI Echo[6]/1998	1.4 (0.8-2.3)	.16
Aronow et al[7]/1998	1.8 (1.2-2.7)	.003
Aronow et al[7]/1998 (LV hypertrophy)	2.8 (1.8-4.4)	.0001
Aronow et al[14]/1989 (LV hypertrophy)	6.56 (NR)	< .01
SPAF[21]/1992	2.6 (1.4-4.9)	.003
SPAF I-III Aspirin[3]/1999 (using 2D echocardiography)	NR	NS
SPAF I-III Aspirin[3]/1999 (using M-mode echocardiography)	NR	NS
Stöllberger et al[4]/1998	NR	NS

2D = two dimensional; LV = left ventricular. See Table S1 and S2 legends for expansion of other abbreviations.

Table S7—*Studies Assessing Valvular Heart Disease[a] as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
AFI Echo[6]/1998 (mitral valve prolapse)	0.29 (NR)	NS
AFI Echo[6]/1998 (mitral valve regurgitation)	1.07 (NR)	NS
Aronow et al[7]/1998 (aortic stenosis)	NR	NS
Aronow et al[7]/1998 (mitral annular calcifications)	NR	NS
Nakagami et al[20]/1998 (mitral valve regurgitation)	0.45 (0.20-0.97)	< .05
van Staa et al[12]/2011	1.65 (1.01-2.71)	< .05

See Table S1 and S2 legends for expansion of abbreviations.
[a]Specific type of valvular heart disease in parentheses.

 DOI: 10.1378/chest.11-2304

Table S8—*Studies Assessing Female Sex as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
AFI[2]/1994	NR	NS
AFI Echo[6]/1998[a]	NR	NS
Aronow et al[7]/1998	NR	NS
Aronow et al[14]/1989	0.85 (NR)	NS
Cabin et al[15]/1990	NR	.014
Fang et al[22]/2005 (all women)	1.6 (1.3-1.9)	< .05
Fang et al[22]/2005 (women aged > 75 vs men aged > 75 y)	1.8 (1.4-2.3)	< .05
Fang et al[22]/2005 (women aged ≤ 75 vs men aged ≤ 75 y)	1.6 (1.0-2.3)	NS
Hart et al[8]/2000 (intermittent AF)	NR	NS
Hart et al[8]/2000 (sustained AF)	1.8	.004
Inoue and Atarashi[19]/2000[b] (paroxysmal AF)	0.50 (0.27-0.93)	.0291
Moulton[13]/1991	NR	NS
Nakagami et al[20]/1998	0.98 (0.55-1.72)	NS
Petersen et al[23]/1990	NR	NS
SPAF III[18]/1998	NR	NS
SPAF I-III Aspirin[3]/1999 (all women)	1.6 (NR)	.01
SPAF I-III Aspirin[3]/1999 (women aged > 75 y)	3.0 (NR)	.002
Stöllberger et al[4]/1998	NR	NS
van Latum et al[11]/1995	1.5 (1.0-2.4)	.05
Van Staa et al[12]/2011[b]	1.05 (0.94-1.19)	NS
Wang et al[5]/2003	1.9 (1.2-3.1)	< .05

See Table S1 and S2 legends for expansion of abbreviations.
[a]Subset of patients from AFI[2] with echocardiographic data available.
[b]Reference sex in original study was female; reported values of RR have been inverted to be directly comparable with other studies.

Table S9—*Studies Assessing Estrogen-Based HRT as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
Fang et al[22]/2005 (22% HRT use)	1.0 (0.7-1.4)	NS
SPAF I-III Aspirin[3]/1999 (33% HRT use at entry)	3.1 (NR)	.007

HRT = hormone replacement therapy. See Table S1 and S2 legends for expansion of other abbreviations.

Table S10—*Studies Assessing Coronary Artery Disease as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
AFI[2]/1994 (history angina)	NR	NS
AFI[2]/1994 (history myocardial infarction)	1.7 (NR)	NS
Aronow et al[14]/1989 (history myocardial infarction)	4.84 (NR)	< .01
Petersen et al[23]/1990 (history myocardial infarction)	1.7 (NR)	.0375
SPAF III[18]/1998	NR	NS
SPAF I-III Aspirin[3]/1999	NR	NS
Stöllberger et al[4]/1998	NR	NS

Other studies have examined peripheral artery disease as an independent predictor of stroke in patients with AF.[24,25] See Table S1 and S2 legends for expansion of abbreviations.

Table S11—*Studies Assessing Carotid Artery Stenosis as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
SPAF II[26]/1994	1.8 (0.5-3.6)	.55

See Table S1 legend for expansion of abbreviations.

Table S12—*Studies Assessing Left Atrial Thrombus as an Independent Predictor of Stroke in Patients With AF*

Study/Year	Multivariate OR/RR (95% CI)	*P* Value
Stöllberger[4]/1998	2.4 (0.9-6.9)	.09

See Table S1 for expansion of abbreviations.
[a]One-half of the 10 patients with left atrial thrombus received anticoagulant therapy, likely blunting the predictive value.

 DOI: 10.1378/chest.11-2304

Table S13—*Stroke Risk Stratification Schema*

Study/Year	Low Risk	Intermediate Risk	High Risk
AFI[3]/1994	Age < 65 y and no risk factors	Age > 65 y and no other risk factors	Prior stroke/TIA, hypertension, diabetes
SPAF[18]/1998	No risk factors	Hypertension, diabetes	Prior stroke/TIA, women > 75 y, men > 75 y with hypertension
$CHADS_2$[27]/2001	Score 0	Score 1-2	Score 3-6
AFI[28]/2003	No risk factors	No intermediate risk category	Previous stroke/TIA, hypertension, diabetes, angina, previous MI
Framingham[5]/2003	Score 0-7	Score 8-15	Score 16-31
NICE guidelines[29]/2006	Age < 65 y with no moderate/high risk factors	Age ≥ 65 y with no high risk factors Age < 75 y with hypertension, diabetes or vascular disease[a]	Previous stroke/TIA or thromboembolic event Age ≥ 75 y with hypertension, diabetes or vascular disease Clinical evidence of valve disease or heart failure, or impaired left ventricular function
ACC/AHA/ESC guidelines[30]/2006	No risk factors	Age ≥ 75 y, or hypertension, or heart failure, or LVEF < 35%, or diabetes	Previous stroke, TIA or embolism, or ≥ 2 moderate risk factors of (age ≥ 75y, hypertension, heart failure, LVEF ≤ 35%, diabetes)
8th ACCP guidelines[31]/2008	No risk factors	Age > 75 y, or hypertension, or moderately or severely impaired LVEF and/or heart failure, or diabetes	Previous stroke, TIA or embolism, or ≥ 2 moderate risk factors of (age ≥ 75y, hypertension, moderately or severely impaired LVEF and/or heart failure, diabetes)
$CHADS_2$-VASc[32]/2009	No risk factors	One combination risk factor: (heart failure/LVEF ≤ 40, hypertension, diabetes, vascular disease,[a] female gender, age 65-74)	Previous stroke, TIA or embolism, or age ≥ 75 y, or ≥ 2 combination risk factors (heart failure/LVEF ≤ 40, hypertension, diabetes, vascular disease,[a] female gender, age 65-74)

ACC/AHA/ESC = American College of Cardiology/American Heart Association/European Society of Cardiology; $CHADS_2$ = congestive heart failure, hypertension, age > 75 years, diabetes, prior stroke or transient ischemic attack [doubled]; LVEF = left ventricular ejection fraction; NICE = National Institute for Health and Clinical Excellence; TIA = transient ischemic attack.
[a]Myocardial infarction, peripheral artery disease, or aortic placque.

References

1. Stroke Risk in Atrial Fibrillation Working Group. Independent predictors of stroke in patients with atrial fibrillation: a systematic review. *Neurology*. 2007;69(6):546-554.
2. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. *Arch Intern Med*. 1994;154(13):1449-1457.
3. Hart RG, Pearce LA, McBride R, Rothbart RM, Asinger RW; The Stroke Prevention in Atrial Fibrillation (SPAF) Investigators. Factors associated with ischemic stroke during aspirin therapy in atrial fibrillation: analysis of 2012 participants in the SPAF I-III clinical trials. *Stroke*. 1999;30(6):1223-1229.
4. Stöllberger C, Chnupa P, Kronik G, et al. Transesophageal echocardiography to assess embolic risk in patients with atrial fibrillation. ELAT Study Group. Embolism in Left Atrial Thrombi. *Ann Intern Med*. 1998;128(8):630-638.
5. Wang TJ, Massaro JM, Levy D, et al. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA*. 2003;290(8):1049-1056.
6. Echocardiographic predictors of stroke in patients with atrial fibrillation: a prospective study of 1066 patients from 3 clinical trials. *Arch Intern Med*. 1998;158(12):1316-1320.
7. Aronow WS, Ahn C, Kronzon I, Gutstein H. Risk factors for new thromboembolic stroke in patients > or = 62 years of age with chronic atrial fibrillation. *Am J Cardiol*. 1998;82(1):119-121.
8. Hart RG, Pearce LA, Rothbart RM, McAnulty JH, Asinger RW, Halperin JL; Stroke Prevention in Atrial Fibrillation Investigators. Stroke with intermittent atrial fibrillation: incidence and predictors during aspirin therapy. *J Am Coll Cardiol*. 2000;35(1):183-187.
9. The Stroke Prevention in Atrial Fibrillation Investigators. Predictors of thromboembolism in atrial fibrillation: I. Clinical features of patients at risk. *Ann Intern Med*. 1992;116(1):1-5.
10. Stöllberger C, Chnupa P, Abzieher C, et al. Mortality and rate of stroke or embolism in atrial fibrillation during long-term follow-up in the embolism in left atrial thrombi (ELAT) study. *Clin Cardiol*. 2004;27(1):40-46.
11. van Latum JC, Koudstaal PJ, Venables GS, van Gijn J, Kappelle LJ, Algra A; European Atrial Fibrillation Trial (EAFT) Study Group. Predictors of major vascular events in patients with a transient ischemic attack or minor ischemic stroke and with nonrheumatic atrial fibrillation. *Stroke*. 1995;26(5):801-806.
12. Van Staa TP, Setakis E, Di Tanna GL, Lane DA, Lip GY. A comparison of risk stratification schemes for stroke in 79,884 atrial fibrillation patients in general practice. *J Thromb Haemost*. 2011;9(1):39-48.
13. Moulton AW, Singer DE, Haas JS. Risk factors for stroke in patients with nonrheumatic atrial fibrillation: a case-control study. *Am J Med*. 1991;91(2):156-161.

 DOI: 10.1378/chest.11-2304

14. Aronow WS, Gutstein H, Hsieh FY. Risk factors for thromboembolic stroke in elderly patients with chronic atrial fibrillation. *Am J Cardiol*. 1989;63(5):366-367.
15. Cabin HS, Clubb KS, Hall C, Perlmutter RA, Feinstein AR. Risk for systemic embolization of atrial fibrillation without mitral stenosis. *Am J Cardiol*. 1990;65(16):1112-1116.
16. Petersen P, Kastrup J, Helweg-Larsen S, Boysen G, Godtfredsen J. Risk factors for thromboembolic complications in chronic atrial fibrillation. The Copenhagen AFASAK study. *Arch Intern Med*. 1990;150(4):819-821.
17. Seidl K, Hauer B, Schwick NG, Zellner D, Zahn R, Senges J. Risk of thromboembolic events in patients with atrial flutter. *Am J Cardiol*. 1998;82(5):580-583.
18. The SPAF III Writing Committee for the Stroke Prevention in Atrial Fibrillation Investigators. Patients with nonvalvular atrial fibrillation at low risk of stroke during treatment with aspirin: Stroke Prevention in Atrial Fibrillation III Study. *JAMA*. 1998;279(16):1273-1277.
19. Inoue H, Atarashi H; Research Group for Antiarrhythmic Drug Therapy. Risk factors for thromboembolism in patients with paroxysmal atrial fibrillation. *Am J Cardiol*. 2000;86(8):852-855.
20. Nakagami H, Yamamoto K, Ikeda U, Mitsuhashi T, Goto T, Shimada K. Mitral regurgitation reduces the risk of stroke in patients with nonrheumatic atrial fibrillation. *Am Heart J*. 1998;136(3):528-532.
21. The Stroke Prevention in Atrial Fibrillation Investigators. Predictors of thromboembolism in atrial fibrillation: II. Echocardiographic features of patients at risk. *Ann Intern Med*. 1992;116(1):6-12.
22. Fang MC, Singer DE, Chang Y, et al. Gender differences in the risk of ischemic stroke and peripheral embolism in atrial fibrillation: the AnTicoagulation and Risk factors In Atrial fibrillation (ATRIA) study. *Circulation*. 2005;112(12):1687-1691.
23. Petersen P. Thromboembolic complications in atrial fibrillation. *Stroke*. 1990;21(1):4-13.
24. Lin LY, Lee CH, Yu CC, et al. Risk factors and incidence of ischemic stroke in Taiwanese with nonvalvular atrial fibrillation - a nation wide database analysis. *Atherosclerosis*. 2011;217:292-295.
25. Frost L, Engholm G, Johnsen S, Møller H, Husted S. Incident stroke after discharge from the hospital with a diagnosis of atrial fibrillation. *Am J Med*. 2000;108(1):36-40.
26. Warfarin versus aspirin for prevention of thromboembolism in atrial fibrillation: Stroke Prevention in Atrial Fibrillation II Study. *Lancet*. 1994;343(8899):687-691.
27. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA*. 2001;285(22): 2864-2870.
28. van Walraven C, Hart RG, Wells GA, et al. A clinical prediction rule to identify patients with atrial fibrillation and a low risk for stroke while taking aspirin. *Arch Intern Med*. 2003;163(8):936-943.
29. NICE. *Atrial Fibrillation: National Clinical Guideline for Management in Primary and Secondary Care*. London, England: Royal College of Physicians' National Collaborating Centre for Chronic Conditions on behalf of the National Institute for Health and Clinical Excellence; 2006.
30. Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation): developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Circulation*. 2006;114(7):e257-354.
31. Singer DE, Albers GW, Dalen JE, et al. Antithrombotic therapy in atrial fibrillation: American College of Chest Physicians evidence-based clinical practice guidelines (8th Edition). *Chest*. 2008;133(Suppl 6):546S-592S.
32. Lip GY, Nieuwlaat R, Pisters R, et al. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the Euro Heart Survey on atrial fibrillation. *Chest*. 2010;137(2):263-272.

 DOI: 10.1378/chest.11-2304

Substance Abuse, 33:373–377, 2012

ISSN: 0889-7077 print / 1547-0164 online
DOI: 10.1080/08897077.2012.662421

Routledge
Taylor & Francis Group

Behavior, Knowledge, and Attitudes Towards Khat Among Yemeni Medical Students and Effects of a Seminar

Paul H. Yi, BA
John S. Kim, BA, MA
Khalil I. Hussein, BA
Richard Saitz, MD, MPH

ABSTRACT. This study describes khat behavior, knowledge, and attitudes among Yemeni medical students (MS) and the effects of a seminar. The students completed a survey ($n = 62$); a subgroup participated in a discussion-based seminar and follow-up survey ($n = 18$). Although the students demonstrated knowledge about khat's health effects and considered it unacceptable for health professionals to chew khat, they disagreed that health professionals should advise patients to quit. Knowledge and attitudes improved post-seminar (not significant, except for a borderline significant increase in students correctly identifying khat as addictive; $P = 0.063$). Although effects were small, seminars may help health professionals address khat use in Yemen.

KEYWORDS. Attitudes, khat, knowledge, medical students, seminar

INTRODUCTION

Khat (also spelled "qat") use is prevalent in Yemen and in many other parts of the Middle East and eastern Africa, especially Somalia and Ethiopia, and it has significant health and socioeconomic consequences (1). Khat is a plant whose active ingredient is cathinone, an amphetamine-like alkaloid with addictive properties that produces a pleasurable stimulant effect (1, 2). Although khat use is not prevalent in Western countries, synthetic cathinone use has risen in the past 4 years in Europe and in the past year in the United States (3). Khat's role in causing

Paul H. Yi, John S. Kim, Khalil I. Hussein, and Richard Saitz are affiliated with the Boston University School of Medicine, Boston, Massachusetts, USA. Richard Saitz is also affiliated with the Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, and Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, USA.

Address correspondence to: Paul H. Yi, BA, 711 S. Ashland Ave., Unit P, Chicago, IL 60607, USA (E-mail: paulyi88@gmail.com).

This article was presented in abstract form at John McCahan Medical Education Day at Boston University School of Medicine on June 8, 2011.

The authors thank Park Street Church Boston for their administrative help and vision to improve global medical education and for helping finance our travel to Yemen; William G. Pearson for his guidance and leadership in organizing our trip; Bridges Intercultural Studies and MERCYemen for seminar curriculum design and logistics; Dr. Rafael Martinez for seminar logistics and survey translation accuracy check; Mohamed A. Kamel and Dr. Abdulmaged Traish for survey translation and accuracy checks; Arthur Celestin and Natalie Witek for helping with survey administration; and our Yemeni colleagues at Taiz University for their time and openness with us.

cardiovascular, oral, hepatic, neurobehavioral, and psychiatric illness is well documented (4). Estimates of daily khat chewing rates in Yemen are as high as 90% for men and 73% for women (1) and, in a country predicted to be the first in the world to run out of water, 40% of Yemen's water supply is directed towards khat irrigation (2).

To our knowledge, there has been no study of knowledge and attitudes towards khat focusing on health professionals who could intervene in Yemen or elsewhere. The recent surge in synthetic cathinone use in Europe and the United States (3) may make khat and cathinone a significant health issue in the western world in the years ahead.

The purpose of this descriptive report is twofold: (1) To describe the behavior, knowledge, and attitudes of medical students at a medical school in Yemen towards khat chewing and (2) to describe these measures in these students before and after a khat seminar conducted through English conversation.

METHODS

Population/Sample

We conducted a written survey of Yemeni medical students at Taiz University (TU) Faculty of Medicine in July 2010. A subgroup participated in one 3-hour seminar on Khat as part of a public health course, open to all TU medical students, entitled "Health Education Through English Conversation Course" (HETECC), and was surveyed a second time at the end of the course. The HETECC course and our survey were advertised by fliers placed around the TU campus and by word of mouth. The course required an English proficiency test.

Survey Administration

We administered the 10-minute survey before and 10 to 11 days after the khat seminar and during the HETECC course. At the beginning of each administration, we described our study to the students in English and informed them that the study would be anonymous, voluntary, and have no impact on their academic standing. Additionally, each survey came attached with a cover letter written in Arabic stating the purpose and the voluntary, anonymous nature of the study. The investigators were not present when the survey was completed and the participants had the option to not participate and to omit answering any survey questions. Participants were not compensated for survey completion.

Seminar

The stated goal of the khat seminar was "to clarify the medical, social, economic, and environmental effects of qat on Yemeni society." There were 3 seminar groups based on English proficiency level, each facilitated by native English speakers in groups of two. The seminar consisted of 6 activities conducted in English: (1) open discussion on an article, titled "A Medical View of Qat," from the *Yemen Times* newspaper; (2) fill-in-the-blank exercise for an article missing words, titled "Qat's Enormous Cost," also from *Yemen Times*; (3) oral reports and discussion of the results of students' informal surveys of friends' and families' weekly khat consumption and expenses, motivation for chewing khat, and willingness to give up khat; (4) discussion of the effects of khat on physical health, the environment, Yemen's economy, and Yemeni society; (5) role-play about 2 common scenarios in Yemeni society involving khat; and (6) discussion of the agricultural economics of khat.

Survey Content and Design

We developed a survey based on the World Health Organization Global Health Professional Survey (GHPS), a standardized questionnaire assessing behavior, knowledge, and attitudes of health care professional students towards tobacco; we modified the version of the GHPS used by Hodgetts et al. to be specific to khat (5). Our survey consisted of 35 questions covering 4 domains: (1) demographics; (2) khat and other substance use; (3) knowledge about khat; and (4) attitudes towards the role and use of khat in Yemeni society. We developed the survey and accompanying cover letter first in English; an Arabic speaker fluent in English subsequently

translated them into Arabic. Two Arabic speakers fluent in English then reviewed the translated survey and assured that the translation was accurate. The Boston University Medical Campus Institutional Review Board and the HETECC course director at TU approved all research activities. Our survey is available upon request.

Analyses

Analyses were descriptive and we used Mann-Whitney, Fisher exact, Wilcoxon signed-rank, and McNemar tests as appropriate.

RESULTS

A total of 62 students completed the survey; the mean age was 22 years, 74% were female, 97% were unmarried, 39% reported ever chewing khat (73% of males and 26% of females), 34% had drunk 8 or more cups of coffee in the previous 30 days, 3% reported ever smoking 100 or more cigarettes, and 3% reported knowing medical students who consume alcohol. Correct response rate to knowledge questions were 90% for khat containing cathinone, 79% for excessive khat use leading to psychosis, 60% for khat chewing (KC) being a risk factor for heart attack, and 90% for up to 90% of Yemeni men chewing khat 3 to 4 hours daily. Regarding attitudes, 87% agreed or strongly agreed that KC is harmful to health, 30% that khat is not addictive, 93% that most people in Yemen chew khat, 6% that khat is not a problem in Yemen, 97% that children and adolescents should not chew khat, 12% that khat chewers are immoral people, 51% that KC is not an immoral activity, 58% that KC makes people more popular with their friends, 27% that KC makes talking to people easier, 47% that KC does not help people to relax, 77% that KC helps people to stay awake, 87% that health professionals (HPs) serve as role models for patients and the public, 8% that it is acceptable for HPs to chew khat, 10% that HPs should routinely ask patients about their KC habits, 94% that HPs should not routinely advise patients to quit KC, 92% that HPs should get specific education on khat, and 50% that they had received education about khat in medical school prior to the survey.

Among khat chewers, 55% consumed 1 or more wraps per day, 29% had chewed khat on 20 or more days in the 30 days prior to the survey, 81% had chewed khat within the previous 7 months, 65% were ready to quit KC, and 20% were thinking about quitting KC. Coffee consumption did not differ between chewers and nonchewers, but chewers were more likely to have used greater than 100 cigarettes in a lifetime and more likely to know a medical student who had used alcohol in the previous 30 days (9.1% of chewers vs. 0% of nonchewers for both items), both borderline significant at $P = 0.064$ and $P = 0.061$, respectively. For all knowledge and attitude survey items, results were not significantly different between chewers and nonchewers, except for borderline significance at $P = 0.069$ towards acceptability of HPs to chew khat (17% of chewers vs. 3% of nonchewers agreed or strongly agreed that it is acceptable).

Of 23 students who participated in the discussion-based khat seminar, 18 completed an identical follow-up survey 10 to 11 days after the seminar. Although in general, knowledge and attitudes improved after the seminar, improvements were not statistically significant except for a borderline improvement in the number of students correctly identifying khat as addictive (33.3% of students pre-seminar vs. 61.1% post-seminar; $P = 0.063$).

DISCUSSION

Less than 50% of Yemeni medical students ever chewed khat, but the majority of males reported ever chewing khat. Of the chewers, two thirds expressed readiness to quit chewing khat. Overall, the students demonstrated good knowledge about khat chewing. Most students also agreed that khat chewing is harmful to health and a problem in Yemen, and that it is not acceptable for health professionals to chew khat. However, paradoxically, they did not agree that health professionals should inquire about or advise their patients to quit their khat chewing habits. Not surprisingly, more khat chewers than nonchewers agreed that it is acceptable for health professionals to chew khat. A seminar had modest effects on knowledge and attitudes towards khat.

The majority of adults and medical students in Yemen chew khat (1, 6), with higher use in males than females (2). Although we found that a minority of students chews khat, when stratified by gender, our results are more consistent with prior estimates. Yemeni citizens have knowledge regarding the physiological properties of khat (2) and Yemeni citizens and Yemeni expatriates in the United Kingdom report attitudes that khat is harmful to health and causes socioeconomic problems (2, 7), which is consistent with our results. The majority of medical students in over 48 countries agree that health professionals serve as role models and that they should advise patients to quit smoking cigarettes (8), which is consistent with our results regarding the students' views of health professionals as role models, but inconsistent with their attitudes towards advising patients to quit chewing khat. A 3-hour lecture-and-discussion–based seminar on tobacco performed at a medical school in Hong Kong was shown to be effective in improving medical students' knowledge of and attitudes towards tobacco control (9), which is not inconsistent with our findings regarding the khat seminar, although we found no statistically significant improvements.

The importance of khat in Yemeni culture and society (1, 2) may help explain why the medical students in our sample were reluctant to agree that health professionals should inquire about or advise their patients to quit khat chewing, even though they agreed that it is not acceptable for health professionals to chew khat themselves; to denounce khat, perhaps, is to denounce Yemeni culture. Another explanation may be in the use of khat for stress relief and academic enhancement within the medical profession. A study of Yemeni physicians concluded that *not* chewing khat was a predictor of burnout (10). In Saudi Arabia, Yemen's geographical neighbor to the north and where khat is prohibited, medical students smoke tobacco to overcome school-related stresses, despite being knowledgeable of its detrimental health effects (11). As the majority of the khat chewers in our sample agreed that khat chewing helps people to stay awake and to relax, khat chewing among Yemeni medical students can thus be considered a cultural analogue to tobacco use in Saudi Arabian medical students to relieve stress and coffee consumption in American medical students to aid in studying.

Our study should be viewed in light of its limitations. One limitation was that our sample was a small convenience sample. Additionally, our seminar subgroup demonstrated high levels of knowledge prior to the seminar, which made it difficult to detect a difference in knowledge after the seminar (ceiling effect).

Our study also had strengths. First, we are unaware of any other study assessing the knowledge, use, and attitudes of Yemeni medical students towards khat. In developing survey items, we took cultural and religious sensitivities into account. We asked about knowledge of alcohol use among medical students, rather than personal alcohol use, because alcohol is illegal in Yemen; however, when asking about additional stimulant use, we failed to take into consideration that tea is the caffeinated beverage of choice over coffee. Third, our survey was anonymous and voluntary and we had no control over the students' grades or standing at the university.

The solution to the paradox of students considering it unacceptable for health professionals to chew khat, but being unwilling to advise patients to quit, is not purely a matter of knowledge, as the students demonstrated good knowledge about khat. Khat's role in Yemeni culture may be a primary reason for these conflicting attitudes. Therefore, perhaps Yemeni public health officials need to make public statements on khat's health hazards before current and future health professionals in Yemen will feel that it is acceptable to get involved with their patients' chewing habits. Studies on khat policy have called for governments to actively pursue regulation of khat focused on harm and risk reduction (12). In this regard, the Yemeni government has made efforts in the past, such as limitations of khat use to certain days of the week (12) and former President Ali Abdullah Saleh has discouraged khat use on several occasions (1). Nevertheless, continued and increased efforts may be critical to empowering current and future health professionals in Yemen to advise their patients about khat. Another part of the solution may involve the clinical education of health professionals. Given khat's unique cultural place in Yemeni society, current and future health professionals

may benefit from being trained to ask about khat in culturally and socially appropriate ways, for example, to have an open discussion about khat and to leave the decision up to the patient. Although our seminar had few effects, the borderline significant effect observed, the likely need for a discussion forum to adequately address the problem, and the effects of such seminars on tobacco suggest that discussion-based seminars may be a part of the solution.

REFERENCES

1. Al-Mugahed L. Khat chewing in Yemen: turning over a new leaf. *Bull World Health Organ*. 2008;86:741–742.

2. Wedegaertner F, al-Warith H, Hillemacher T, et al. Motives for khat use and abstinence in Yemen—a gender perspective. *BMC Public Health*. 2010;10:735.

3. The Drug Enforcement Administration. Request for information on synthetic cathinones. *Microgram Bull*. 2011;44:31–32.

4. Al-Motarreb A, Al-Habori M, Broadley KJ. Khat chewing, cardiovascular diseases and other internal medical problems: the current situation and directions for future research. *J Ethnopharmacol*. 2010;132:540–548.

5. Hodgetts G, Broers T, Godwin M. Smoking behaviour, knowledge and attitudes among Family Medicine physicians and nurses in Bosnia and Herzegovina. *BMC Fam Pract*. 2004;5:12.

6. Laswar AK, Darwish H. Prevalence of cigarette smoking and khat chewing among Aden University medical students and their relationship to BP and body mass index. *Saudi J Kidney Dis Transpl*. 2009;20:862–866.

7. Kassim S, Croucher R. Khat chewing amongst UK resident male Yemeni adults: an exploratory study. *Int Dent J*. 2006;56:97–101.

8. Warren C, Sinha D, Lee J, Lea V, Jones N. Tobacco use, exposure to secondhand smoke, and cessation counseling among medical students: cross-country data from the Global Health Professions Student Survey (GHPSS), 2005–2008. *BMC Public Health*. 2011;11:72.

9. Chung TW, Lam TH, Cheng YH. Knowledge and attitudes about smoking in medical students before and after a tobacco seminar. *Med Educ*. 1996;30:290–295.

10. Al-Dubai SAR, Rampal KG. Prevalence and Associated factors of burnout among doctors in Yemen. *J Occup Health*. 2010;52:58–65.

11. Al-Haqwi AI, Tamim H, Asery A. Knowledge, attitude and practice of tobacco smoking by medical students in Riyadh, Saudi Arabia. *Ann Thorac Med*. 2010;5:145–148.

12. Klein A, Beckerleg S, Hailu D. Regulating khat—dilemmas and opportunities for the international drug control system. *Int J Drug Policy*. 2009;20:509–513.

Hindawi Publishing Corporation
Pain Research and Treatment
Volume 2012, Article ID 167062, 7 pages
doi:10.1155/2012/167062

Research Article

Attitudinal Barriers to Analgesic Use among Patients with Substance Use Disorders

Leah Zallman,[1] Sonia L. Rubens,[2] Richard Saitz,[2,3] Jeffrey H. Samet,[2,4] Christine Lloyd-Travaglini,[5] and Jane Liebschutz[2,4]

[1] *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA*
[2] *Department of Medicine, Cambridge Health Alliance, Cambridge, MA 02139, USA*
[3] *Department of Epidemiology, Boston University School of Public Health, Boston, MA 02118, USA*
[4] *Department of Community Health Sciences, Boston University School of Public Health, Boston, MA 02118, USA*
[5] *Data Coordinating Center, Boston University School of Public Health, Boston, MA 02118, USA*

Correspondence should be addressed to Leah Zallman, lzallman@challiance.org

Received 14 November 2011; Revised 29 February 2012; Accepted 2 March 2012

Academic Editor: Howard Smith

Attitudinal barriers towards analgesic use among primary care patients with chronic pain and substance use disorders (SUDs) are not well understood. We evaluated the prevalence of moderate to significant attitudinal barriers to analgesic use among 597 primary care patients with chronic pain and current analgesic use with 3 subscales from the Barriers Questionaire II: concern about side effects, fear of addiction, and worry about reporting pain to physicians. Concern about side effects was a greater barrier for those with opioid use disorders (OUDs) and non-opioid SUDs than for those with no SUD (OR (95% CI): 2.30 (1.44–3.68), $P < 0.001$ and 1.64 (1.02–2.65), $P = 0.041$, resp.). Fear of addiction was a greater barrier for those with OUDs as compared to those with non-opioid SUDs (OR (95% CI): 2.12 (1.04–4.30), $P = 0.038$) and no SUD (OR (95% CI): 2.69 (1.44–5.03), $P = 0.002$). Conversely, participants with non-opioid SUDs reported lower levels of worry about reporting pain to physicians than those with no SUD (OR (95% CI): 0.43 (0.24–0.76), $P = 0.004$). Participants with OUDs reported higher levels of worry about reporting pain than those with non-opioid SUDs (OR (95% CI): 1.91 (1.01–3.60), $P = 0.045$). Concerns about side effects and fear of addiction can be barriers to analgesic use, moreso for people with SUDs and OUDs.

1. Introduction

Chronic pain is commonly encountered in primary care [1], where the majority of patients with substance use disorders (SUDs) receive health care [2]. Patients with SUDs are at significant risk of pain [3, 4] and are likely to be undertreated for pain. Not only do patients with SUDs report continued pain despite engaging in medical care [5], but few providers caring for these patients follow recommended guidelines [6]. Many pathways may be responsible for such undertreatment, from health systems issues, like insurance coverage and limited access to specialists, to clinician attitudes and skills [7]. For example, clinicians may be appropriately reluctant to prescribe opioid pain medications to patients with histories of substance use disorders out of concern that patients may divert or misuse them [8]. One possible and relatively unexplored reason for undertreatment is patient attitudinal barriers to analgesic medications, such as concerns about side effects, fear of addiction, and worries about reporting pain to physicians. Indeed, patients with SUDs who are HIV positive or are enrolled in methadone maintenance programs report significant attitudinal barriers to pain medications, including concerns regarding addiction potential, need for escalating doses, and difficulty communicating with their clinicians regarding pain [9, 10]. However, such attitudinal barriers to analgesic use, such as opioid medication use, among patients with SUDs in primary care settings have not been well described.

Several patient characteristics have been associated with attitudes that may result in more barriers to the use of analgesic medications (i.e., higher attitudinal barriers) including non-White race, lower education, more physical symptoms

[10], older age [11], higher pain severity and disability [12], unemployment [13], and depression [14], while the data on gender are mixed [13, 15]. Other characteristics have not been explored but seem likely to impact attitudinal barriers, such as whether patients are currently or previously addicted to substances, the substances to which they may be addicted (e.g., opioids versus nonopioid drugs), and recent use of analgesics. For example, patients with past addiction, who are currently in recovery, may have more negative attitudes about opioid analgesics than those currently addicted. Similarly, patients with opioid use disorders (OUDs) may be more likely than patients with nonopioid SUDs to have negative attitudes regarding the addiction potential of opioid pain medications. Finally, though psychiatric illnesses, such as depression and anxiety, are known to be associated with more attitudinal barriers [14], the ways in which other psychiatric illnesses, like posttraumatic stress disorder, common among people with chronic pain [16, 17] and SUDs [18], affect attitudinal barriers has not yet been explored.

Our objective was to compare attitudinal barriers to analgesic use among patients with SUDs and those without, across three domains: concern about side effects, fear of addiction, and worries about reporting pain to physicians. We hypothesized that the majority of primary care patients with SUDs would have moderate-to-severe attitudinal barriers and that those with SUDs would have greater barriers in all three domains as compared to those without SUDs. We also explored whether two subcategories of SUDs—(1) current versus lifetime disorder and (2) OUD versus nonopioid SUD—report greater barriers. We hypothesized that people with lifetime SUDs as compared to current SUDs, and OUDs as compared to nonopioid SUDs, would have greater barriers in all three domains.

2. Materials and Methods

2.1. Participants and Setting. We conducted a cross-sectional study of a sample of primary care patients with chronic pain. We recruited participants between February 2005 and August 2006 from waiting rooms of an academic, urban hospital primary care practice. We approached patients waiting for appointments and included patients who were 18–60 years of age, spoke English, had pain for three months or more, reported use of any analgesic medication (over-the-counter or prescription) in the prior month, and had a scheduled primary care appointment. Informed consent was obtained from eligible patients. Trained interviewers administered surveys. We compensated participants with $10. The Boston University Medical Center Institutional Review Board approved the study, and a certificate of confidentiality was obtained from the National Institutes of Health.

2.2. Instruments

Dependent Variables. The dependent variables of interest were the Barriers Questionnaire II (BQ-II) subscale scores. The BQ-II, a validated 27-item self-report survey, was originally designed to measure barriers to obtaining pain relief for cancer patients [11]. We selected the BQ-II because it is a widely used instrument and has been adapted for use with adolescent cancer patients [19] and HIV/AIDS populations [10], and is used in ambulatory settings [10, 19]. We eliminated all questions referring to cancer pain and the immune system in order to make it applicable to a primary care population. The resulting 9-item version, displayed in Table 1, contains three subscales including concern about side effects, fear of addiction, and worries about reporting pain to physicians. The questionnaire refers to "pain medications" and does not specify opioid analgesics. The items are scaled on a 6-point Likert scale from 0 (do not agree) to 5 (agree very much). A-one point increase on the overall BQ-II has previously been shown to be associated with a twofold greater likelihood of being undermedicated for pain control [10]. A BQ subscale score cutoff of ≥ 3 has been previously used in the literature [10]; we considered a score of ≥ 3 to be indicative of moderate-to-significant barriers.

Independent Variables. The main independent variables of interest were mutually exclusive categories of substance use disorders identified using the Composite International Diagnostic Interview v. 2.1 Drug Use Disorder module, a well-validated interview instrument that yields current and past DSM-IV diagnoses [20, 21].

We assigned participants into three mutually exclusive categories of substance use disorders: those with "OUDs" (which included opioid abuse and dependence), "nonopioid SUDs" (met criteria for a substance use disorder but not an OUD—includes abuse and dependence on alcohol or other drug) and no SUD. If a participant met criteria for more than one substance use disorder and one was opioids, we categorized them as "OUD."

Covariates. We adjusted analyses for factors known to be associated with attitudinal barriers to analgesic use [11, 13–15], or, in the absence of data, due to clinical suspicion that a factor might be associated with attitudinal barriers. These factors were (1) sociodemographic variables including sex, race/ethnicity (Black/African American, Hispanic/Latino/Other, White), employment (unemployed, including receiving disability payments versus employed part or full time), education (<high school, high school+); (2) depression (major or other versus none) determined by the Patient Health Questionnaire (PHQ) for Depression, a 9-item validated measure correlated with past two-week major depression [22]; (3) somatic symptom severity (high versus others) determined by the Patient Health Questionnaire-15, a validated measure that correlates with somatization disorder [23]; (4) a lifetime history of posttraumatic stress disorder (PTSD) diagnosis derived from the CIDI v. 2.1 PTSD module [20]; (5) opioid prescription in the past year. We determined the proportion of participants meeting Prescription Drug Use Disorder (PDUD), defined as any opiate use disorder that was not heroin alone, but did not include PDUD as a covariate in the models because it is a subset of OUD. We determined the proportion of participants with moderate versus severe pain and disability using the Graded

Table 1: Modified Barriers Questionnaire.

Subscale	Item
Side effects	Drowsiness from pain medicine is difficult to control.
	When you use pain medicine, your body becomes used to its effects and pretty soon it will not work anymore.
	Using pain medicine blocks your ability to know if you have any new pain.
	Pain medicine can keep you from knowing what is going on in your body.
Addiction	There is a danger of becoming addicted to pain medicine.
	Pain medicine is very addictive.
Reporting pain to physicians	It is important to be strong by not talking about pain.
	It is important for the doctor to focus on curing illness, and not waste time controlling pain.
	Doctors might find it annoying to be told about pain.

Chronic Pain Scale [24], a 10-item validated measure of pain and disability; we did not include pain severity and disability as a covariate in the models because there is little variation in this measure (90% reported severe pain).

2.3. Statistical Analysis

Prevalence of Moderate-to-Significant Barriers. We compared the proportion of participants with moderate-to-significant barriers in each category of SUD using chi square analyses.

Current versus Lifetime SUD. We conducted a sensitivity analysis to examine the relationship between the nonmutually exclusive categories of lifetime (includes both current and past) and current SUDs, and prevalence of moderate-to-significant barriers. Because we did not detect large differences, we used lifetime disorders only for subsequent analyses.

OUD versus Nonopioid SUD. We conducted bivariate analyses using general linear models with the Duncan multiple range test for differences to examine the relationship between the independent variable of interest (OUD, nonopioid SUD and no SUD), and the BQ-II subscale scores. Because the three categories differed in their relationship to BQ-II subscale scores (data not shown), we maintained this variable as a three-part variable in further analyses.

Logistic Regression Analysis of Barriers. In bivariate analyses, the main independent variable impacted the subscales in opposite directions. We therefore examined the subscales separately. We first conducted bivariate analyses to examine the relationship between covariates and the barrier subscales. We then constructed three multivariable logistic regression models—one for each subscale—in order to examine the relationship between substance use category (OUD, nonopioid SUD, and no SUD) and the odds of having moderate-to-significant barriers. Each model was adjusted for the main independent variable and all covariates. Tests for colinearity among the covariates were carried out using correlation coefficients and any colinear variables were eliminated.

3. Results

3.1. Participant Characteristics. Of 822 eligible patients, 597 (73%) enrolled and completed the research interview. Enrollees were more likely than those who declined enrollment to be black (61% versus 55%, $P = 0.04$), less likely to take over-the-counter pain medication (66% versus 79%, $P < 0.001$), and more likely to take opioid pain medication (41% versus 30%, $P = 0.002$).

Participant characteristics are shown in Table 2. A majority of the study sample was female, unemployed, and nonwhite. Roughly half of participants had no SUD, one quarter had an OUD and one quarter had a nonopioid SUD. Of those with an OUD, 104/138 (75%) met criteria for a lifetime prescription drug use disorder. As noted in previous literature [25], the prevalence of OUD is higher among whites (28%) versus blacks (19%) and among males (32%) versus females (17%).

3.2. Attitudinal Barriers

Prevalence of Moderate-to-Significant Attitudinal Barriers. In the full sample, moderate-to-significant fear of side effects (60%) and concern about addiction (71%) were common while worry about reporting pain to physicians was relatively uncommon (27%). Participants with OUDs more often reported moderate-to-significant fears of addiction than those with no nonopioid SUDs (89% versus 78%, $P = 0.02$) and no SUDs (89% versus 75%, $P = 0.01$). Participants with nonopioid SUDs less often reported moderate-to-significant worries about reporting pain to physicians than those with no SUDs (18% versus 31%, $P = 0.01$).

Current Versus Lifetime SUD. There were no differences in prevalence of moderate-to-significant barriers between those with current and lifetime SUD for any of the subscales (concern about side effects (66% versus 64%), fear of addiction (83% versus 78%), and worries about reporting pain to physicians (23% versus 18%)).

Unadjusted Analyses. Unadjusted logistic regression analyses of moderate-to-significant barriers comparing substance use groups are presented in Table 3. Concern about side effects and fear of addiction scores were greater among participants with OUDs than they were for those with no SUDs. Similarly, participants with OUDs had greater fear of addiction than those with nonopioid SUDs. Conversely, scores reflecting worries about reporting pain to physicians

Table 2: Demographic and clinical characteristics of a sample of primary care patients, stratified by substance abuse ($N = 597$).

Variable	Group			P value
	OUD	Nonopioid SUD	No SUD	
	$n = 138$ n (%)	$n = 118$ n (%)	$n = 341$ n (%)	
Age, mean in years (SD)	45 (8.9)	45 (8.1)	46 (10.4)	0.66
Race/ethnicity				<0.001[1,2]
Black/African American	70 (51%)	71 (61%)	222 (65%)	
Hispanic/Latino/Other	25 (18%)	23 (20%)	81 (24%)	
White	43 (31%)	23 (20%)	37 (11%)	
Gender				<0.001[1,2]
Female	59 (43%)	54 (46%)	237 (70%)	
Male	79 (57%)	64 (54%)	104 (31%)	
Employment status				<0.001[1,2]
Unemployed or disabled	97 (70%)	83 (70%)	181 (53%)	
Full-/part-time	41 (30%)	35 (30%)	160 (47%)	
Education				0.56
Less than high school	35 (25%)	37 (31%)	94 (28%)	
High school or above	103 (75%)	81 (69%)	247 (72%)	
Depression				0.03[1]
Major and/or other	68 (49%)	54 (46%)	127 (37%)	
None	70 (51%)	64 (54%)	214 (63%)	
Pain severity and disability				0.01[2]
Severe	127 (92%)	113 (96%)	295 (87%)	
Moderate	11 (8%)	5 (4%)	46 (13%)	
Somatic symptom severity				0.06
High	54 (39%)	48 (41%)	104 (31%)	
Low/medium	84 (61%)	70 (59%)	237 (70%)	
PTSD				<0.001[1,2]
Lifetime history	63 (46%)	56 (47%)	100 (29%)	
No history	75 (54%)	62 (53%)	241 (71%)	
Opioid prescription (past year)				0.65
Yes	60 (44%)	46 (40%)	132 (39%)	
No	77 (56%)	68 (60%)	205 (61%)	

[1] Significance < 0.05 for comparison between OUD and no SUD.
[2] Significance < 0.05 for comparison between nonopioid SUD and no SUD.

Table 3: Logistic regression models of moderate-to-significant barriers stratified by substance use disorder[†*] ($N = 597$).

	Side Effect		Addiction		Reporting Pain	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Unadjusted						
OUD versus No SUD	**2.09 (1.37, 3.20)**	**<0.001**	**2.78 (1.54, 5.00)**	**<0.001**	0.83 (0.53, 1.29)	0.41
Nonopioid SUD versus no SUD	1.49 (0.97, 2.29)	0.07	1.20 (0.73, 1.97)	0.48	**0.49 (0.29, 0.83)**	**0.01**
OUD versus nonopioid SUD	1.40 (0.83, 2.37)	0.21	**2.32 (1.16, 4.62)**	**0.02**	1.69 (0.93, 3.10)	0.09
Adjusted						
OUD versus no SUD	**2.30 (1.44, 3.68)**	**<0.001**	**2.69 (1.44, 5.03)**	**0.002**	0.82 (0.50, 1.35)	0.44
Non opioid SUD versus no SUD	**1.64 (1.02, 2.65)**	**0.04**	1.27 (0.74, 2.19)	0.39	**0.43 (0.24, 0.76)**	**0.004**
OUD versus nonopioid SUD	1.40 (0.81, 2.43)	0.23	**2.12 (1.04, 4.30)**	**0.04**	**1.91 (1.01, 3.60)**	**0.045**

[†] Moderate-to-significant barriers defined as score of ≥3 on subscale (range 0–5).
[*] Models adjusted for gender, employment, depression, somatic symptom severity, education, race, PTSD, and recent opioid use.

were lower for those with nonopioid SUDs than for those with no SUDs.

Adjusted Analyses. Multivariate logistic regression analyses of moderate-to-significant attitudinal barrier subscale scores among participants with OUDs, nonopioid SUDs, and no SUD are also shown in Table 3. The greater concerns about side effects and fear of addiction among participants with OUDs compared with those with no SUD persisted in adjusted analyses. Similarly, those with OUDs continued to report higher fears of addiction and higher concern about reporting pain as compared to those with nonopioid SUDs. Differences in worry about reporting pain to physicians between those with nonopioid SUDs remained lower than among those with no SUDs in adjusted analyses. After adjustment, participants with nonopioid SUDs had higher concerns about side effects than those with no SUDs, and participants with OUDs had higher concerns about reporting pain than those with nonopioid SUDs.

4. Discussion

In our sample of 597 primary care patients with chronic pain, moderate-to-significant fear about side effects and concern about addiction was common, and more so for participants with OUDs and nonopioid SUDs than those with no SUDs. Conversely, moderate-to-significant worries about reporting pain to physicians were less common overall, and less so for participants with nonopioid SUDs. In adjusted analyses, participants with OUDs reported more concern about side effects and fear of addiction yet similar worries about reporting pain to physicians as those with no SUD. Participants with nonopioid SUDs reported lower worries about reporting pain to physicians than those with no SUD and those with OUDs. Interestingly, those with current and lifetime substance use disorders reported similar barriers with regard to all three subscales.

Why might people with OUDs and chronic pain report a greater degree of attitudinal barriers to analgesic use than participants with nonopioid SUDs? In our study, there were no differences in baseline characteristics between those with OUDs and those with nonopioid SUDs, which suggests that there is something about the experience of the opioid itself (such as the potential addictive quality of opioids) that is associated with increased barriers. Because relapse or worse addiction may be triggered by exposure to opioid medications, those with OUDs may have heightened concern compared to those with nonopioid SUDs, who may perceive less risk of addiction with exposure to opioids. This finding is consistent with other studies showing that HIV-infected patients with a prior injection drug use history are more concerned than their noninjection drug using counterparts about the addictive potential of pain medications [10].

Greater attitudinal barriers among participants with OUDs are particularly interesting because this observation is in contradiction to the experience of many clinicians who describe patients with pain and opioid abuse as "drug seeking" [26]. In fact, while the majority of participants in our study reported attitudinal barriers to analgesic medications, in a study of physicians' perceptions of barriers to AIDS pain management, only a minority of physicians (24%) believed patient reluctance to take opioids to be a barrier to pain management [27]. This discrepancy between physician perception of patient reluctance to take opioids and patient concerns about taking opioids may be a source of miscommunication between patients and physicians. In fact, a clinical encounter for a patient with a substance use history that is focused on pain may have competing demands on the part of both the clinician and the patient. Clinicians balance ambiguity in pain assessment, treatment goals, and treatment risks [28]; similarly, though attitudinal barriers may play a role for the patient, these barriers may not be expressed due to a desire to obtain pain medications for pain relief or misuse, including for their euphoric effect or diversion, all of which are common among patients with prior substance use disorders [5].

The finding that attitudinal barriers were similar for those with current and lifetime opioid use disorders suggests that patients with past SUDs may harbor as many negative attitudes regarding concerns about side effects and fear of addiction potential as patients currently struggling with opioid addiction. This caution about taking pain medications in patients with past SUDs and chronic pain may be appropriate. While opioids may provide good short-term pain relief, chronic opioid use can produce a number of problems, including relapse [29]. However, these perceptions may also lead to undertreatment of pain.

Why might individuals with SUDs have lesser attitudinal barriers to reporting pain to doctors than individuals without SUDs? Patients in this study all suffered from chronic pain and may therefore have developed a certain level of comfort discussing pain. In addition, patients in our study also had ongoing relationships with regular primary care clinicians, a key characteristic identified by patients as essential to good communication and clinical decision making [30]. Finally, patients with SUDs may find it easier to talk about pain than other areas, such as addiction or mental health.

Several limitations apply to our findings. The sample was derived from a single academic urban hospital primary care practice and is therefore potentially subject to idiosyncratic practices, which may limit its generalizability. Nonetheless, the sample reflects a demographically heterogeneous group and is likely similar to patients in many other urban primary care practices in the US. In addition, the questions focused on possible side effects, addiction potential, and reporting of pain, and not about attitudes towards requesting or receiving prescriptions for opioid pain medications. Thus, we may have missed the potentially contentious part of the medical encounter: namely, the dispensation of opioid prescriptions. Finally, the questions did not explicitly delineate which classes of pain medications (e.g., opioid versus nonopioid analgesics) to consider. We expect that most participants assumed questions were referring to opioid medications; indeed recent opioid prescription was common in our sample, suggesting that many participants had experience with prescription opioids.

Despite these limitations, the study offers a preliminary look into the nature of patient attitudinal barriers to analgesic use in primary care patients with pain and addictions. It also offers an alternative perspective on patient-physician interactions than that previously described. Prior studies have demonstrated that the patient-physician interaction regarding pain medications is characterized by mutual mistrust and difficult communication [26]. Primary care providers display discomfort and avoidance in discussing unhealthy alcohol use [31], suggesting that physicians experience difficulty communicating about substance use. However, in this study, patients with SUDs reported low attitudinal barriers to reporting pain to clinicians. They also reported similar concerns to those of physicians, namely, concern regarding pain medication side effects and addiction potential of these medications [10]. These findings may therefore highlight common ground between patients and physicians that offers a basis for patient-physician alliance rather than the previously described "mutual mistrust" [26].

5. Conclusions

In summary, patients with opioid use disorders and chronic pain have greater attitudinal barriers to analgesic use than those without opioid use disorders. Patients with substance use disorders have lower attitudinal barriers to reporting pain to physicians. Given the substantial functional, social, and psychological consequences of undertreated pain, there is an imperative for physicians to work with patients who have comorbid pain and substance use disorders to address their attitudinal barriers to pain medications. This study highlights a modifiable barrier to effectively treating chronic pain in patients with SUD. Further clinical care and research should focus on addressing attitudinal barriers so that patients may have maximum pain relief without compromising recovery from addiction.

Disclouse

Portions of this work were presented at the Society of General Internal Medicine New England Regional (March 13, 2009, Boston, MA) and National (May 14, 2009, Miami, FL) meetings.

Acknowledgments

This work was supported by K23 DA016665 from the National Institute of Drug Abuse, K24 AA015674 from the National Institute on Alcohol Abuse and Alcoholism, and an Institutional National Research Service Award no. T32HP12706 from the Health Services and Resources Administration to support the Harvard Medical School Fellowship in General Medicine and Primary Care. There are no financial conflicts of interest. This work was presented at the Society of General Internal Medicine National Meeting in Miami, FL in May, 2009 and the New England Regional Society of General Internal Medicine in Boston, MA in March, 2009.

References

[1] B. Breuer, R. Cruciani, and R. K. Portenoy, "Pain management by primary care physicians, pain physicians, chiropractors, and acupuncturists: a national survey," *Southern Medical Journal*, vol. 103, no. 8, pp. 738–747, 2010.

[2] D. B. Kamerow, H. A. Pincus, and D. I. Macdonald, "Alcohol abuse, other drug abuse, and mental disorders in medical practice. Prevalence, costs, recognition, and treatment," *JAMA*, vol. 255, no. 15, pp. 2054–2057, 1986.

[3] A. Rosenblum, H. Joseph, C. Fong, S. Kipnis, C. Cleland, and R. K. Portenoy, "Prevalence and characteristics of chronic pain among chemically dependent patients in methadone maintenance and residential treatment facilities," *JAMA*, vol. 289, no. 18, pp. 2370–2378, 2003.

[4] R. N. Jamison, J. Kauffman, and N. P. Katz, "Characteristics of methadone maintenance patients with chronic pain," *Journal of Pain and Symptom Management*, vol. 19, no. 1, pp. 53–62, 2000.

[5] L. Hansen, J. Penko, D. Guzman, D. R. Bangsberg, C. Miaskowski, and M. B. Kushel, "Aberrant behaviors with prescription opioids and problem drug use history in a community-based cohort of HIV-infected individuals," *Journal of Pain and Symptom Management*, vol. 42, no. 6, pp. 893–902, 2011.

[6] P. J. Lum, S. Little, M. Botsko et al., "Opioid-prescribing practices and provider confidence recognizing opioid analgesic abuse in HIV primary care settings," *Journal of Acquired Immune Deficiency Syndromes*, vol. 56, supplement 1, pp. S91–S97, 2011.

[7] D. T. Barry, K. S. Irwin, E. S. Jones et al., "Opioids, chronic pain, and addiction in primary care," *Journal of Pain*, vol. 11, no. 12, pp. 1442–1450, 2010.

[8] B. Bhamb, D. Brown, J. Hariharan, J. Anderson, S. Balousek, and M. F. Fleming, "Survey of select practice behaviors by primary care physicians on the use of opioids for chronic pain," *Current Medical Research and Opinion*, vol. 22, no. 9, pp. 1859–1865, 2006.

[9] A. Karasz, L. Zallman, K. Berg, M. Gourevitch, P. Selwyn, and J. Arnstein, "The experience of chronic severe pain in patients undergoing methadone maintenance treatment," *Journal of Pain and Symptom Management*, vol. 28, no. 5, pp. 517–525, 2004.

[10] W. Breitbart, S. Passik, M. V. McDonald et al., "Patient-related barriers to pain management in ambulatory AIDS patients," *Pain*, vol. 76, no. 1-2, pp. 9–16, 1998.

[11] S. Gunnarsdottir, H. S. Donovan, R. C. Serlin, C. Voge, and S. Ward, "Patient-related barriers to pain management: the barriers questionnaire II (BQ-II)," *Pain*, vol. 99, no. 3, pp. 385–396, 2002.

[12] V. T. Potter, C. E. Wiseman, S. M. Dunn, and F. M. Boyle, "Patient barriers to optimal cancer pain control," *Psycho-Oncology*, vol. 12, no. 2, pp. 153–160, 2003.

[13] R. Dawson, D. E. Sellers, J. A. Spross, E. S. Jablonski, D. R. Hoyer, and M. Z. Solomon, "Do patients' beliefs act as barriers to effective pain management behaviors and outcomes in patients with cancer-related or noncancer-related pain?" *Oncology Nursing Forum*, vol. 32, no. 2, pp. 363–374, 2005.

[14] J. Edrington, A. Sun, C. Wong et al., "Barriers to pain management in a community sample of Chinese American patients with cancer," *Journal of Pain and Symptom Management*, vol. 37, no. 4, pp. 665–675, 2009.

[15] S. E. Ward, N. Goldberg, V. Miller-McCauley et al., "Patient-related barriers to management of cancer pain," *Pain*, vol. 52, no. 3, pp. 319–324, 1993.

[16] C. L. Villano, A. Rosenblum, S. Magura, C. Fong, C. Cleland, and T. F. Betzler, "Prevalence and correlates of posttraumatic stress disorder and chronic severe pain in psychiatric outpatients," *Journal of Rehabilitation Research and Development*, vol. 44, no. 2, pp. 167–177, 2007.

[17] G. J. G. Asmundson and J. Katz, "Understanding the co-occurrence of anxiety disorders and chronic pain: state-of-the-art," *Depression and Anxiety*, vol. 26, no. 10, pp. 888–901, 2009.

[18] K. A. Villagonzalo, S. Dodd, F. Ng, S. Mihaly, A. Langbein, and M. Berk, "The relationship between substance use and posttraumatic stress disorder in a methadone maintenance treatment program," *Comprehensive Psychiatry*, vol. 52, no. 5, pp. 562–566, 2011.

[19] S. Ameringer, R. C. Serlin, S. H. Hughes, S. A. Frierdich, and S. Ward, "Concerns about pain management among adolescents with cancer: developing the adolescent barriers questionnaire," *Journal of Pediatric Oncology Nursing*, vol. 23, no. 4, pp. 220–232, 2006.

[20] World Health Organization, *Composite International Diagnostic Interview (CIDI): Version 2.1*, World Health Organization, 1997.

[21] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, American Psychiatric Association, Washington, DC, USA, 4th edition, 1994.

[22] R. L. Spitzer, K. Kroenke, and J. B. Williams, "Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire," *JAMA*, vol. 282, no. 18, pp. 1737–1744, 1999.

[23] K. Kroenke, R. L. Spitzer, and J. B. W. Williams, "The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms," *Psychosomatic Medicine*, vol. 64, no. 2, pp. 258–266, 2002.

[24] M. Von Korff, J. Ormel, F. J. Keefe, and S. F. Dworkin, "Grading the severity of chronic pain," *Pain*, vol. 50, no. 2, pp. 133–149, 1992.

[25] L. T. Wu, G. E Woody, C. Yang, P. Mannelli, and D. G. Blazer, "Differences in onset and abuse/dependence episodes between prescription opioids and heroin: results from the National Epidemiologic Survey on Alcohol and Related Conditions," *Journal of Substance Abuse and Rehabilitation*, no. 2, pp. 77–88, 2011.

[26] J. O. Merrill, L. A. Rhodes, R. A. Deyo, G. A. Marlatt, and K. A. Bradley, "Mutual mistrust in the medical care of drug users: the keys to the "Narc" cabinet," *Journal of General Internal Medicine*, vol. 17, no. 5, pp. 327–333, 2002.

[27] W. Breitbart, M. Kaim, and B. Rosenfeld, "Clinicians' perceptions of barriers to pain management in AIDS," *Journal of Pain and Symptom Management*, vol. 18, no. 3, pp. 203–212, 1999.

[28] K. M. Berg, J. H. Arnsten, G. Sacajiu, and A. Karasz, "Providers' experiences treating chronic pain among opioid-dependent drug users," *Journal of General Internal Medicine*, vol. 24, no. 4, pp. 482–488, 2009.

[29] J. Højsted and P. Sjøgren, "Addiction to opioids in chronic pain patients: a literature review," *European Journal of Pain*, vol. 11, no. 5, pp. 490–518, 2007.

[30] E. Murray, L. Pollack, M. White, and B. Lo, "Clinical decision-making: patients' preferences and experiences," *Patient Education and Counseling*, vol. 65, no. 2, pp. 189–196, 2007.

[31] K. A. McCormick, N. E. Cochran, A. L. Back, J. O. Merrill, E. C. Williams, and K. A. Bradley, "How primary care providers talk to patients about alcohol: a qualitative study," *Journal of General Internal Medicine*, vol. 21, no. 9, pp. 966–972, 2006.